Accuracy of the Bethe approximation for hyperparameter estimation in probabilistic image processing



2004 J. Phys. A: Math. Gen. 37 8675

(http://iopscience.iop.org/0305-4470/37/36/007)

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

J. Phys. A: Math. Gen. **37** (2004) 8675–8695 PII: S0305-4470(04)76248-4

# Accuracy of the Bethe approximation for hyperparameter estimation in probabilistic image processing

**Kazuyuki Tanaka**[1], **Hayaru Shouno**[2], **Masato Okada**[3,4] **and D M Titterington**[5]

[1] Graduate School of Information Sciences, Tohoku University, Aramaki-aza-aoba 09, Aoba-ku, Sendai 980-8579, Japan
[2] Department of Computer Science and Systems Engineering, Yamaguchi University, Tokiwadai 2-16-1, Ube 755-8611, Japan
[3] RIKEN Brain Science Institute, Hirosawa 2-1, Wakou 351-0456, Japan
[4] Graduate School of Frontier Sciences, The University of Tokyo, Kashiwanoha 5-1-5, Kashiwa 277-8562, Japan
[5] Department of Statistics, University of Glasgow, Glasgow G12 8QQ, UK

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

Received 16 February 2004, in final form 13 May 2004
Published 24 August 2004
Online at stacks.iop.org/JPhysA/37/8675
doi:10.1088/0305-4470/37/36/007

**Abstract**

We investigate the accuracy of statistical-mechanical approximations for the estimation of hyperparameters from observable data in probabilistic image processing, which is based on Bayesian statistics and maximum likelihood estimation. Hyperparameters in statistical science correspond to interactions or external fields in the statistical-mechanics context. In this paper, hyperparameters in the probabilistic model are determined so as to maximize a marginal likelihood. A practical algorithm is described for grey-level image restoration based on a Gaussian graphical model and the Bethe approximation. The algorithm corresponds to loopy belief propagation in artificial intelligence. We examine the accuracy of hyperparameter estimation when we use the Bethe approximation. It is well known that a practical algorithm for probabilistic image processing can be prescribed analytically when a Gaussian graphical model is adopted as a prior probabilistic model in Bayes' formula. We are therefore able to compare, in a numerical study, results obtained through mean-field-type approximations with those based on exact calculation.

PACS numbers: 02.50−r, 02.50.Cw, 02.50.Tt, 05.20.−y, 05.50.+q, 75.10.Nr, 89.70.+c

  

## 1. Introduction

Some statistical-mechanical models and techniques are useful in information processing [1]. In particular, many physicists and computer scientists are interested in the application of advanced mean-field methods to create new iterative algorithms for solving certain practical information processing problems [2]. In the area of artificial intelligence, belief propagation has been investigated as a tool for probabilistic inference [3, 4]. The mathematical structure of belief propagation is basically same as the transfer matrix method for classical spin systems on lattices with tree structures in statistical mechanics [5]. In statistics and computer science, the classical spin systems are interpretable as *graphical models*. In some important problems in computer science, we have to treat graphical models on lattices with loops. Some computer scientists have also applied belief propagation to graphical models on lattices with loops and have proposed some approximate algorithms [6, 7]. They refer to this as *loopy belief propagation*. Some statistical physicists have pointed out that the loopy belief propagation is equivalent to the Bethe approximation in statistical mechanics [8, 2]. Yedidia *et al* [9] proposed an update rule for generalized belief propagation and showed that the rule is equivalent to the cluster variation method in statistical mechanics.

Freeman *et al* [10] applied generalized belief propagation to some computer vision approaches. Many approaches that use generalized belief propagation are based on probabilistic models involving Markov random fields and Bayes' formula [11, 12]. It is well known that probabilistic image processing is useful for many practical computer vision problems [13, 14]. The probabilistic models are equivalent to classical spin systems with short-range interactions and spatially non-uniform external fields [15]. Thus, advanced mean-field methods, including the Bethe approximation, are very useful in probabilistic image processing based on Bayes' formula and Markov random fields.

Weiss and Freeman [7] investigated the accuracy of the Bethe approximation in probabilistic image restoration when a Gaussian graphical model is adopted as the *a priori* probabilistic model. In this case, we can derive an exact expression for the restored image by using the multi-dimensional Gauss integral formula and the discrete Fourier transformation [15, 16, 20]. In [7], Weiss and Freeman investigated the average and the variance of the random variable of the light intensity at each pixel in computer vision and concluded that the accuracy of the average is good but that of the variance is not adequate if the Bethe approximation is used.

In probabilistic image restoration, an *a posteriori* probability distribution and a marginal likelihood are constructed based on Bayesian statistics and maximum marginal likelihood estimation. The *a posteriori* probability distribution corresponds to a classical spin system in statistical mechanics. The input is a degraded image and the output is both the corresponding restored image and a set of estimates of hyperparameters which correspond to interaction parameters in the classical spin systems. The hyperparameters can be determined so as to maximize a marginal likelihood [21, 22]. The restored image is determined from the *a posteriori* probability distribution based on the estimated set of hyperparameters. The implication of the results in [7] is that, although for fixed values of the hyperparameters the restored images obtained by the exact expression and the Bethe approximation are very similar, the accuracy of the marginal likelihood obtained by the Bethe approximation may not be good enough.

Our main purpose in the present paper is to assess the accuracy of the estimates obtained from the maximization of marginal likelihood by means of the Bethe approximation in grey-level image restorations. In order to fulfil the above goal, we mainly discuss a solvable probabilistic model, for which there is exact analytical treatment, rather than more realistic

models. Such solvable models are often unrealistically simple. However, by comparing the results obtained using approximations with those from the exact calculations, we can assess the accuracy of the approximate methods explicitly through numerical experiments. In the present paper, we consider grey-level image restoration, formulated by means of the Bayesian approach, and we adopt additive white Gaussian noise and a Gaussian graphical model as the degradation process and the *a priori* probability model respectively. It is assumed that the intensity at each pixel can take any real value. We compare maximum marginal likelihood estimation based on the Bethe approximation with the versions based on the mean-field approximation and on the exact expression for the marginal likelihood that is available in this scenario.

The present paper is organized as follows. In section 2, we summarize the basic framework of Bayesian image restoration and maximization of the marginal likelihood for hyperparameter estimation based on a Gaussian graphical model. In section 3, we give the schemes for the calculation of the marginal likelihood if we use the mean-field approximation and the Bethe approximation. Section 4 describes numerical experiments and section 5 provides concluding remarks.

## 2. Probabilistic image processing based on a gaussian graphical model

In computer vision, images are typically defined on a set of points arranged on a rectangular lattice. Each point is called a *pixel*. At each pixel, the intensity of light is represented as an integer or a real number. We consider an image on a rectangular lattice $\Omega \equiv \{(x, y) | x = 1, 2, \ldots, M, y = 1, 2, \ldots, N\}$ such that the intensity at each pixel takes a real value in the range $(-\infty, +\infty)$. Here, $x$ and $y$ denote the spatial coordinates of pixels. A monochrome digital image is then expressed as a two-dimensional light intensity function $f_{x,y}$, where $f_{x,y}$ is proportional to the brightness of the image at the point $(x, y)$. The rectangular lattice $\Omega$ is assumed to have periodic boundary conditions in both $x$- and $y$-directions. The intensities at pixel $(x, y)$ in the original image and the degraded image are regarded as random variables denoted by $F_{x,y}$ and $G_{x,y}$, respectively, and the random fields of intensities in the original image and the observed, degraded image are represented by $\boldsymbol{F} \equiv \{F_{x,y} | (x, y) \in \Omega\}$ and $\boldsymbol{G} \equiv \{G_{x,y} | (x, y) \in \Omega\}$, respectively. The actual original image and degraded image are denoted by $\boldsymbol{f} = \{f_{x,y}\}$ and $\boldsymbol{g} = \{g_{x,y}\}$, respectively.

The probability density for the original image $\boldsymbol{f}$, $\mathcal{P}(\boldsymbol{F} = \boldsymbol{f})$, is called the *a priori* probability density of the image. As a consequence of Bayes' formula, the *a posteriori* probability density $\mathcal{P}(\boldsymbol{F} = \boldsymbol{f} | \boldsymbol{G} = \boldsymbol{g})$, that the original image is $\boldsymbol{f}$ when the given degraded image is $\boldsymbol{g}$, is expressed as

$$\mathcal{P}(\boldsymbol{F} = \boldsymbol{f} | \boldsymbol{G} = \boldsymbol{g}) = \frac{\mathcal{P}(\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}) \mathcal{P}(\boldsymbol{F} = \boldsymbol{f})}{\int \mathcal{P}(\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{z}) \mathcal{P}(\boldsymbol{F} = \boldsymbol{z}) \, \mathrm{d}\boldsymbol{z}}, \tag{1}$$

where $\int \mathrm{d}\boldsymbol{z} \equiv \prod_{(x,y) \in \Omega} \int_{-\infty}^{+\infty} \mathrm{d}z_{x,y}$. The probability $\mathcal{P}(\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f})$ is the conditional probability density that the degraded image is $\boldsymbol{g}$ when the original image is $\boldsymbol{f}$, and it describes the degradation process.

In the present paper, it is assumed that the degraded image $\boldsymbol{g}$ is generated from the original image $\boldsymbol{f}$ by the addition of white Gaussian noise with mean 0 and variance $\sigma^2$, so that

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

where $|\Omega|$ is the total number of pixels[6]. Moreover, the *a priori* probability density that the original image is $\boldsymbol{f}$ is assumed to be

$$\mathcal{P}(\boldsymbol{F}=\boldsymbol{f}|\alpha)\equiv\frac{1}{\mathcal{Z}_{\mathrm{PR}}(\alpha)}\prod_{(x,y)\in\Omega}\exp\left(-\frac{1}{2}\alpha(f_{x,y}-f_{x+1,y})^2-\frac{1}{2}\alpha(f_{x,y}-f_{x,y+1})^2\right), \tag{3}$$

which represents local spatial correlation among the pixel intensities. The denominator $\mathcal{Z}_{\mathrm{PR}}(\alpha)$ in equation (3) is a normalization constant which is defined by

$$\mathcal{Z}_{\mathrm{PR}}(\alpha)\equiv\int\prod_{(x,y)\in\Omega}\exp\left(-\frac{1}{2}\alpha(f_{x,y}-f_{x+1,y})^2-\frac{1}{2}\alpha(f_{x,y}-f_{x,y+1})^2\right)\mathrm{d}\boldsymbol{z}. \tag{4}$$

By substituting equations (2) and (3) into equation (1), we obtain

$$\mathcal{P}(\boldsymbol{F}=\boldsymbol{f}|\boldsymbol{G}=\boldsymbol{g},\alpha,\sigma)=\frac{1}{\mathcal{Z}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)}\exp(-E(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma)), \tag{5}$$

where

$$E(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma)\equiv\sum_{(x,y)\in\Omega}\left(\frac{1}{2\sigma^2}(f_{x,y}-g_{x,y})^2+\frac{1}{2}\alpha(f_{x,y}-f_{x+1,y})^2+\frac{1}{2}\alpha(f_{x,y}-f_{x,y+1})^2\right), \tag{6}$$

$$\mathcal{Z}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)\equiv\int\exp(-E(\boldsymbol{z}|\boldsymbol{g},\alpha,\sigma))\,\mathrm{d}\boldsymbol{z}. \tag{7}$$

In the maximum likelihood approach, values for the hyperparameters $\alpha$ and $\sigma$ are determined so as to maximize the marginal likelihood $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma)$, where

$$\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma)\equiv\int\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{z},\sigma)\mathcal{P}(\boldsymbol{F}=\boldsymbol{z}|\alpha)\,\mathrm{d}\boldsymbol{z}. \tag{8}$$

We denote these maximizers by $\widehat{\alpha}$ and $\widehat{\sigma}$:

$$(\widehat{\alpha},\widehat{\sigma})=\arg\max_{(\alpha,\sigma)}\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma). \tag{9}$$

Given the estimates $\widehat{\alpha}$ and $\widehat{\sigma}$, the restored image $\widehat{\boldsymbol{f}}=\{\widehat{f}_{x,y}|(x,y)\in\Omega\}$ is determined by

$$\widehat{f}_{x,y}\equiv\int z_{x,y}\mathcal{P}(\boldsymbol{F}=\boldsymbol{z}|\boldsymbol{G}=\boldsymbol{g},\widehat{\alpha},\widehat{\sigma})\,\mathrm{d}\boldsymbol{z}. \tag{10}$$

This way of producing a restored image is called maximum posterior marginal estimation [23, 24]. (Strictly speaking the formula in (10) represents the posterior marginal expectation, but it also provides the 'maximum' because of the Gaussian nature of the posterior.)

The marginal likelihood $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma)$ can be expressed in terms of the free energies $\mathcal{F}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)\equiv-\ln(\mathcal{Z}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma))$ and $\mathcal{F}_{\mathrm{PR}}(\alpha)\equiv-\ln\bigl(\mathcal{Z}_{\mathrm{PR}}(\alpha)\bigr)$ as follows:

$$\begin{aligned}\ln(\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma))&=\ln(\mathcal{Z}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma))-\ln(\mathcal{Z}_{\mathrm{PR}}(\alpha))-|\Omega|\ln(\sqrt{2\pi}\sigma)\\&=-\mathcal{F}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)+\mathcal{F}_{\mathrm{PR}}(\alpha)-|\Omega|\ln(\sqrt{2\pi}\sigma).\end{aligned} \tag{11}$$

If we use the multi-dimensional Gaussian integral formula and the discrete Fourier transformation, the marginal likelihood $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma)$ and the restored image $\widehat{f}_{x,y}$ can be expressed [15–19] as follows,

$$\begin{aligned}\ln(\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma))=&-\frac{|\Omega|}{2}\ln(2\pi)-\frac{1}{2}\sum_{(p,q)\in\Omega}\ln(1+\alpha\sigma^2\gamma(p,q))+\frac{|\Omega|}{2}\ln(\alpha)\\&+\frac{1}{2}\sum_{(p,q)\in\Omega}\ln(\gamma(p,q))-\frac{1}{2}\sum_{(p,q)\in\Omega}|G(p,q)|^2\frac{\alpha\gamma(p,q)}{1+\alpha\sigma^2\gamma(p,q)},\end{aligned} \tag{12}$$

[6] We denote the number of whole elements belonging to a set $A$ by $|A|$.

and

$$\widehat{f}_{x,y} = \frac{1}{\sqrt{|\Omega|}} \sum_{(p,q)\in\Omega} \frac{1}{1+\widehat{\alpha}\,\widehat{\sigma}^2\gamma(p,q)} \left( \cos\left(\frac{2\pi px}{M}+\frac{2\pi qy}{N}\right) \mathrm{Re}(G(p,q)) + \sin\left(\frac{2\pi px}{M}+\frac{2\pi qy}{N}\right) \mathrm{Im}(G(p,q)) \right), \tag{13}$$

where

$$G(p,q) \equiv \frac{1}{\sqrt{|\Omega|}} \sum_{(x,y)\in\Omega} g_{x,y} \exp\left(-\mathrm{i}\frac{2\pi px}{M} - \mathrm{i}\frac{2\pi qy}{N}\right), \tag{14}$$

$$\gamma(p,q) \equiv 4 - 2\cos\left(\frac{2\pi p}{M}\right) - 2\cos\left(\frac{2\pi q}{N}\right). \tag{15}$$

Thus, in the present framework there are closed-form expressions for the marginal likelihood and the restored image, when additive white Gaussian noise and the Gaussian graphical model are adopted as the degradation process and the prior probabilistic model, respectively. By applying the mean-field and Bethe approximations to the calculation of the marginal likelihood $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma)$ and the restored image $\widehat{f}_{x,y}$, we will be able to measure the accuracy of the advanced mean-field methods, relative to the exact treatment.

## 3. Mean-field and Bethe approximations

In this section, we present both the mean-field and the Bethe approximations for probabilistic models of the general form

$$\rho(\boldsymbol{f}) = \frac{1}{\mathcal{Z}} \prod_{(x,y)\in\Omega} \psi_{x,y}(f_{x,y})\phi(f_{x,y}, f_{x+1,y})\phi(f_{x,y}, f_{x,y+1}) \tag{16}$$

where

$$\psi_{x,y}(\xi) \equiv \exp\left(-\tfrac{1}{2}\beta(\xi - g_{x,y})^2\right), \tag{17}$$

$$\phi(\xi,\xi') \equiv \exp\left(-\tfrac{1}{2}\alpha(\xi - \xi')^2\right), \tag{18}$$

and

$$\mathcal{Z} \equiv \int \prod_{(x,y)\in\Omega} (\psi_{x,y}(z_{x,y})\phi(z_{x,y}, z_{x+1,y})\phi(z_{x,y}, z_{x,y+1}))\,\mathrm{d}\boldsymbol{z}. \tag{19}$$

By setting $\beta = 1/\sigma^2$, we obtain the free energy $\mathcal{F}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma) \equiv -\ln\mathcal{Z}_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)$ as $\mathcal{F} \equiv -\ln\mathcal{Z}$. By setting $\beta = 0$, we obtain the free energy $\mathcal{F}_{\mathrm{PR}}(\alpha) \equiv -\ln\mathcal{Z}_{\mathrm{PR}}(\alpha)$ as $\mathcal{F} \equiv -\ln\mathcal{Z}$.

For the purpose of the mean-field approximation, we introduce the average of $f_{x,y}$ with respect to $\rho(\boldsymbol{f})$, defined by

$$m_{x,y} \equiv \int z_{x,y}\rho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}. \tag{20}$$

By setting $\beta = 1/\sigma^2$, we obtain the restored image $\widehat{\boldsymbol{f}}$ as $\{m_{x,y}|(x,y)\in\Omega\}$. Now we assume that, with high probability, the approximate equalities $(f_{x,y}-m_{x,y})(f_{x+1,y}-m_{x+1,y}) \simeq 0$ and $(f_{x,y}-m_{x,y})(f_{x,y+1}-m_{x,y+1}) \simeq 0$ are valid. These equalities can be rewritten as

$$f_{x,y}f_{x+1,y} \simeq m_{x,y}f_{x+1,y} + m_{x+1,y}f_{x,y} - m_{x,y}m_{x+1,y} \tag{21}$$

$$f_{x,y}f_{x,y+1} \simeq m_{x,y}f_{x,y+1} + m_{x,y+1}f_{x,y} - m_{x,y}m_{x,y+1}. \tag{22}$$

By substituting equations (21) and (22) into equations (19) and (20), we obtain

$$m_{x,y} \simeq \frac{\beta g_{x,y} + \alpha \sum_{(x',y') \in \mathbf{c}_{x,y}} m_{x',y'}}{\beta + 4\alpha} \tag{23}$$

and

$$\mathcal{F} = -\ln \mathcal{Z} \simeq -\sum_{(x,y)\in\Omega} \left( \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\beta + 4\alpha) + \frac{\left(\beta g_{x,y} + \alpha \sum_{(x',y') \in \mathbf{c}_{x,y}} m_{x',y'}\right)^2}{2(\beta + 4\alpha)} \right.$$
$$\left. - \frac{1}{2}\beta {g_{x,y}}^2 - \alpha m_{x,y} m_{x+1,y} - \alpha m_{x,y} m_{x,y+1} \right). \tag{24}$$

Here $\mathbf{c}_{x,y} \equiv \{(x \pm 1, y), (x, y \pm 1)\}$ is the set of all the nearest-neighbour pixels of $(x, y)$. In dealing with equations (23) and (24), the free energy $\mathcal{F}$ is obtained by solving the simultaneous fixed-point equations (23) numerically and by substituting the set of solutions $\{m_{x,y} | (x, y) \in \Omega\}$ into the right-hand side of equation (24). The simultaneous fixed-point equations (23) are solved by the following iterative algorithm:

Iterative algorithm for solving the simultaneous fixed-point equations (23)

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

$$a_{x,y}(r) \leftarrow \frac{\beta g_{x,y} + \alpha \sum_{(x',y') \in \mathbf{c}_{x,y}} a_{x',y'}(r-1)}{\beta + 4\alpha} \qquad ((x, y) \in \Omega). \tag{25}$$

*Step 3*. Update $R \leftarrow r$ and $m_{x,y} \leftarrow a_{x,y}(R)$ $((x, y) \in \Omega)$. Stop if

$$\sum_{(x,y)\in\Omega} |a_{x,y}(r) - a_{x,y}(r-1)| < \varepsilon, \tag{26}$$

for prescribed $\varepsilon$, and go to step 2 otherwise.

It is usually adequate to set $\varepsilon = 10^{-6}$. In the denominator of equation (25), the summation $\sum_{(x',y') \in \mathbf{c}_{x,y}} a_{x',y'}(r)$ can be evaluated in $\mathcal{O}(1)$ time per pixel $(x, y)$, because the number of elements in the set $\mathbf{c}_{x,y}$ is equal to 4 per pixel in the present paper. Hence the simultaneous update rules given in equation (25) evaluate $\mathcal{O}(|\Omega|)$ computations so that a total of $\mathcal{O}(|\Omega|)$ computations are required per update.

For the Bethe approximation, we introduce two types of marginal probability densities, defined by

$$\rho_{x,y}(f_{x,y}) \equiv \int \delta(f_{x,y} - z_{x,y}) \rho(\boldsymbol{z})\, \mathrm{d}\boldsymbol{z}, \tag{27}$$

$$\rho_{x,y}^{x',y'}(f_{x,y}, f_{x',y'}) = \rho_{x',y'}^{x,y}(f_{x',y'}, f_{x,y}) \equiv \int \delta(f_{x,y} - z_{x,y}) \delta(f_{x',y'} - z_{x',y'}) \rho(\boldsymbol{z})\, \mathrm{d}\boldsymbol{z}. \tag{28}$$

For consistency, these probability densities should satisfy

$$\int_{-\infty}^{+\infty} \rho_{x,y}(\zeta)\, \mathrm{d}\zeta = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho_{x,y}^{x+1,y}(\zeta, \zeta')\, \mathrm{d}\zeta\, \mathrm{d}\zeta'$$
$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho_{x,y}^{x,y+1}(\zeta, \zeta')\, \mathrm{d}\zeta\, \mathrm{d}\zeta' = 1 \qquad ((x, y) \in \Omega) \tag{29}$$

$$\rho_{x,y}(\xi) = \int_{-\infty}^{+\infty} \rho_{x,y}^{x+1,y}(\xi,\zeta)\,\mathrm{d}\zeta = \int_{-\infty}^{+\infty} \rho_{x,y}^{x,y+1}(\xi,\zeta)\,\mathrm{d}\zeta$$
$$= \int_{-\infty}^{+\infty} \rho_{x-1,y}^{x,y}(\zeta,\xi)\,\mathrm{d}\zeta = \int_{-\infty}^{+\infty} \rho_{x,y-1}^{x,y}(\zeta,\xi)\,\mathrm{d}\zeta \qquad ((x,y)\in\Omega, \xi\in(-\infty,+\infty)). \tag{30}$$

The Bethe free energy for a probabilistic model of the form given in equation (16) is given by

$$\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\middle|(x,y)\in\Omega\right\}\right] \equiv \sum_{(x,y)\in\Omega} \mathcal{F}_{x,y}[\rho_{x,y}]$$
$$+ \sum_{(x,y)\in\Omega} \left(\mathcal{F}_{x,y}^{x+1,y}\left[\rho_{x,y}^{x+1,y}\right] - \mathcal{F}_{x,y}[\rho_{x,y}] - \mathcal{F}_{x+1,y}[\rho_{x+1,y}]\right)$$
$$+ \sum_{(x,y)\in\Omega} \left(\mathcal{F}_{x,y}^{x,y+1}\left[\rho_{x,y}^{x,y+1}\right] - \mathcal{F}_{x,y}[\rho_{x,y}] - \mathcal{F}_{x,y+1}[\rho_{x,y+1}]\right), \tag{31}$$

where

$$\mathcal{F}_{x,y}[\rho_{x,y}] \equiv \int_{-\infty}^{+\infty} \rho_{x,y}(\zeta)\ln\left(\frac{\rho_{x,y}(\zeta)}{\psi_{x,y}(\zeta)}\right)\mathrm{d}\zeta, \tag{32}$$

$$\mathcal{F}_{x,y}^{x',y'}\left[\rho_{x,y}^{x',y'}\right] \equiv \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \rho_{x,y}^{x',y'}(\zeta,\zeta')\ln\left(\frac{\rho_{x,y}^{x',y'}(\zeta,\zeta')}{\psi_{x,y}(\zeta)\phi(\zeta,\zeta')\psi_{x',y'}(\zeta')}\right)\mathrm{d}\zeta\,\mathrm{d}\zeta'. \tag{33}$$

The approximate forms of the marginal probabilities $\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\middle|(x,y)\in\Omega\right\}$ are derived from the extremum conditions of the Bethe free energy $\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\middle|(x,y)\in\Omega\right\}\right]$ with respect to the marginal probability distributions $\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\middle|(x,y)\in\Omega\right\}$ under the constraints given in equations (29) and (30). The marginal probability densities can be given approximately by expressions of the forms

$$\rho_{x,y}(\xi) \simeq \frac{1}{\mathcal{Z}_{x,y}}\psi_{x,y}(\xi)\prod_{(x',y')\in\mathbf{c}_{x,y}} \mathcal{M}_{x,y}^{x',y'}(\xi), \tag{34}$$

$$\rho_{x,y}^{x',y'}(\xi,\xi') \simeq \frac{1}{\mathcal{Z}_{x,y}^{x',y'}}\psi_{x,y}(\xi)\phi(\xi,\xi')\psi_{x',y'}(\xi')$$
$$\times\left(\prod_{(x'',y'')\in\mathbf{c}_{x,y}\setminus(x',y')} \mathcal{M}_{x,y}^{x'',y''}(\xi)\right)\left(\prod_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)} \mathcal{M}_{x',y'}^{x'',y''}(\xi')\right). \tag{35}$$

Here $\mathcal{Z}_{x,y}$ and $\mathcal{Z}_{x,y}^{x',y'}$ are the normalization constants of $\rho_{x,y}(\xi)$ and $\rho_{x,y}^{x',y'}(\xi,\xi')$, respectively. Though these forms may not be so familiar to some physicists, $\ln\left(\mathcal{M}_{x,y}^{x',y'}(\xi)\right)$ corresponds to the effective field from $(x',y')$ to $(x,y)$ in the conventional Bethe approximation. In probabilistic inference, the quantity $\mathcal{M}_{x,y}^{x',y'}(\xi)$ is referred to as a message propagated from $(x',y')$ to $(x,y)$. If we substitute equations (34) and (35) into equation (30), the deterministic equations for $\mathcal{M}_{x,y}^{x',y'}(\xi)$ can be derived as

$$\mathcal{M}_{x,y}^{x',y'}(\xi) = \frac{\int_{-\infty}^{+\infty}\phi(\xi,\xi')\psi_{x',y'}(\xi')\prod_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\mathcal{M}_{x',y'}^{x'',y''}(\xi')\,\mathrm{d}\xi'}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\phi(\xi'',\xi')\psi_{x',y'}(\xi')\prod_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\mathcal{M}_{x',y'}^{x'',y''}(\xi')\,\mathrm{d}\xi'\,\mathrm{d}\xi''}. \tag{36}$$

In the Bethe approximation, the free energy $\mathcal{F} = -\ln\mathcal{Z}$ of the probabilistic model defined by equations (16)–(19) is approximately expressed as

$$\mathcal{F} = -\ln\mathcal{Z} \simeq -\sum_{(x,y)\in\Omega}\ln\mathcal{Z}_{x,y} - \sum_{(x,y)\in\Omega}\left(\ln\mathcal{Z}_{x,y}^{x+1,y} - \ln\mathcal{Z}_{x,y} - \ln\mathcal{Z}_{x+1,y}\right) - \sum_{(x,y)\in\Omega}\left(\ln\mathcal{Z}_{x,y}^{x,y+1} - \ln\mathcal{Z}_{x,y} - \ln\mathcal{Z}_{x,y+1}\right), \tag{37}$$

which is the extreme value of the Bethe free energy $\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x+1,y} \middle| (x,y)\in\Omega\right\}\right]$ with respect to the marginal probability distributions $\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1} \middle| (x,y)\in\Omega\right\}$ under the constraint conditions given in equations (29) and (30). We remark that the Bethe free energy $\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1} \middle| (x,y)\in\Omega\right\}\right]$ does not provide a bound for the true free energy $\mathcal{F} = -\ln\mathcal{Z}$, whereas a mean-field free energy is a bound for the true free energy [2]. In the Bethe free energy $\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1} \middle| (x,y)\in\Omega\right\}\right]$, the first term $\sum_{(x,y)\in\Omega}\mathcal{F}_{x,y}[\rho_{x,y}]$ can be regarded as the contribution of each single pixel. The second term $\sum_{(x,y)\in\Omega}\left(\mathcal{F}_{x,y}^{x+1,y}\left[\rho_{x,y}^{x+1,y}\right] - \mathcal{F}_{x,y}[\rho_{x,y}] - \mathcal{F}_{x+1,y}[\rho_{x+1,y}]\right)$ and the third term $\sum_{(x,y)\in\Omega}\left(\mathcal{F}_{x,y}^{x,y+1}\left[\rho_{x,y}^{x,y+1}\right] - \mathcal{F}_{x,y}[\rho_{x,y}] - \mathcal{F}_{x,y+1}[\rho_{x,y+1}]\right)$ correspond to the contributions of nearest-neighbour pairs of pixels. The right-hand side of equation (31) implies that we are neglecting the other free energies for many-body marginal probability distributions for sets of pixels that are larger than the nearest-neighbour pairs. The mean-field approximation corresponds to the minimization of the approximate free energy $\mathcal{F}_{\mathrm{MF}}[\{\rho_{x,y}|(x,y)\in\Omega\}]$ obtained by setting $\rho_{x,y}^{x',y'}(\xi,\xi') = \rho_{x,y}(\xi)\rho_{x',y'}(\xi')$ in the right-hand side of equation (31). These then are the physical interpretations of the mean-field and Bethe approximations.

Now we assume that $\mathcal{M}_{x,y}^{x',y'}(\xi)$ can be approximately expressed as

$$\mathcal{M}_{x,y}^{x',y'}(\xi) \simeq \sqrt{\frac{\lambda_{x,y}^{x',y'}}{2\pi}}\exp\left(-\frac{\lambda_{x,y}^{x',y'}}{2}\left(\xi - \mu_{x,y}^{x',y'}\right)^2\right). \tag{38}$$

The simultaneous fixed-point equation (36) can be reduced to the following equation:

$$\sqrt{\frac{\lambda_{x,y}^{x',y'}}{2\pi}}\exp\left(-\frac{\lambda_{x,y}^{x',y'}}{2}\left(\xi - \mu_{x,y}^{x',y'}\right)^2\right) = \frac{1}{\sqrt{(2\pi)^2\det\left(\boldsymbol{K}_{x,y}^{x',y'}\right)}}\int_{-\infty}^{+\infty}\exp\left(-\frac{1}{2}\left(\begin{pmatrix}\xi\\ \xi'\end{pmatrix} - \boldsymbol{K}_{x,y}^{x',y'}\boldsymbol{u}_{x,y}^{x',y'}\right)^{\mathrm{T}} \times \left(\boldsymbol{K}_{x,y}^{x',y'}\right)^{-1}\left(\begin{pmatrix}\xi\\ \xi'\end{pmatrix} - \boldsymbol{K}_{x,y}^{x',y'}\boldsymbol{u}_{x,y}^{x',y'}\right)\right)\mathrm{d}\xi', \tag{39}$$

where

$$\boldsymbol{K}_{x,y}^{x',y'} \equiv \begin{pmatrix}\alpha & -\alpha\\ -\alpha & \beta + \alpha + \displaystyle\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\lambda_{x',y'}^{x'',y''}\end{pmatrix}^{-1}, \tag{40}$$

$$\boldsymbol{u}_{x,y}^{x',y'} \equiv \begin{pmatrix}0\\ \beta g_{x',y'} + \displaystyle\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\mu_{x',y'}^{x'',y''}\lambda_{x',y'}^{x'',y''}\end{pmatrix}. \tag{41}$$

Multiplying by $\xi^2$ and $\xi$ and integrating both sides of equation (39) with respect to $\xi$, we can derive the following fixed-point equations for $\left\{\lambda_{x,y}^{x',y'},\ \mu_{x,y}^{x',y'}\,\middle|\,(x',y')\in\mathbf{c}_{x,y},\ (x,y)\in\Omega\right\}$:

$$\frac{1}{\lambda_{x,y}^{x',y'}}=\frac{1}{\alpha}+\frac{1}{\beta+\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\lambda_{x',y'}^{x'',y''}}, \tag{42}$$

$$\mu_{x,y}^{x',y'}=\frac{\beta g_{x',y'}+\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\mu_{x',y'}^{x'',y''}\lambda_{x',y'}^{x'',y''}}{\beta+\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\lambda_{x',y'}^{x'',y''}}. \tag{43}$$

If we substitute equation (38) into equations (34) and (35), the one- and two-body marginal probability densities are approximately rewritten as follows:

$$\rho_{x,y}(\xi)\simeq\sqrt{\frac{\beta+\sum_{(x',y')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x',y'}}{2\pi}}\times\exp\left(-\frac{1}{2}\left(\beta+\sum_{(x',y')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x',y'}\right)\left(\xi-\frac{\beta g_{x,y}+\sum_{(x',y')\in\mathbf{c}_{x,y}}\mu_{x,y}^{x',y'}\lambda_{x,y}^{x',y'}}{\beta+\sum_{(x',y')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x',y'}}\right)^2\right), \tag{44}$$

$$\rho_{x,y}^{x',y'}(\xi,\xi')\simeq\frac{1}{\sqrt{(2\pi)^2\det\left(\boldsymbol{R}_{x,y}^{x',y'}\right)}}\times\exp\left(-\frac{1}{2}\left(\begin{pmatrix}\xi\\ \xi'\end{pmatrix}-\boldsymbol{R}_{x,y}^{x',y'}\boldsymbol{v}_{x,y}^{x',y'}\right)^{\mathrm{T}}\left(\boldsymbol{R}_{x,y}^{x',y'}\right)^{-1}\left(\begin{pmatrix}\xi\\ \xi'\end{pmatrix}-\boldsymbol{R}_{x,y}^{x',y'}\boldsymbol{v}_{x,y}^{x',y'}\right)\right), \tag{45}$$

where

$$\boldsymbol{R}_{x,y}^{x',y'}\equiv\begin{pmatrix}\beta+\alpha+\displaystyle\sum_{(x'',y'')\in\mathbf{c}_{x,y}\setminus(x',y')}\lambda_{x,y}^{x'',y''} & -\alpha\\ -\alpha & \beta+\alpha+\displaystyle\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\lambda_{x',y'}^{x'',y''}\end{pmatrix}^{-1}, \tag{46}$$

$$\boldsymbol{v}_{x,y}^{x',y'}\equiv\begin{pmatrix}\beta g_{x,y}+\displaystyle\sum_{(x'',y'')\in\mathbf{c}_{x,y}\setminus(x',y')}\mu_{x,y}^{x',y'}\lambda_{x,y}^{x',y'}\\ \beta g_{x',y'}+\displaystyle\sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\mu_{x',y'}^{x,y}\lambda_{x',y'}^{x,y}\end{pmatrix}. \tag{47}$$

The mean $m_{x,y}$ defined by equation (20) can be calculated by means of the equation

$$m_{x,y}=\int_{-\infty}^{+\infty}\xi\rho_{x,y}(\xi)\,\mathrm{d}\xi\simeq\frac{\beta g_{x,y}+\sum_{(x',y')\in\mathbf{c}_{x,y}}\mu_{x,y}^{x',y'}\lambda_{x,y}^{x',y'}}{\beta+\sum_{(x',y')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x',y'}}. \tag{48}$$

In terms of $\lambda_{x,y}^{x',y'}$ and $\mu_{x,y}^{x',y'}$, the normalization constants $\mathcal{Z}_{x,y}$ and $\mathcal{Z}_{x,y}^{x',y'}$ in equations (34) and (35) are expressed as

$$\ln\mathcal{Z}_{x,y}=\frac{1}{2}\ln(2\pi)+\frac{1}{2}\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\ln\left(\frac{\lambda_{x,y}^{x'',y''}}{2\pi}\right)-\frac{1}{2}\ln\left(\beta+\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x'',y''}\right)-\frac{1}{2}\beta g_{x,y}{}^2+\frac{\left(\beta g_{x,y}+\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\mu_{x,y}^{x'',y''}\lambda_{x,y}^{x'',y''}\right)^2}{2\left(\beta+\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x'',y''}\right)}-\frac{1}{2}\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\left(\mu_{x,y}^{x'',y''}\right)^2\lambda_{x,y}^{x'',y''}, \tag{49}$$

$$\ln \mathcal{Z}_{x,y}^{x',y'} = \ln(2\pi) + \frac{1}{2} \sum_{(x'',y'') \in \mathbf{c}_{x,y} \setminus (x',y')} \ln \left( \frac{\lambda_{x,y}^{x'',y''}}{2\pi} \right) + \frac{1}{2} \sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} \ln \left( \frac{\lambda_{x',y'}^{x'',y''}}{2\pi} \right)$$
$$+ \frac{1}{2} \ln \left( \det \left( \boldsymbol{R}_{x,y}^{x',y'} \right) \right) + \frac{1}{2} \left( \boldsymbol{v}_{x,y}^{x',y'} \right)^{\mathrm{T}} \boldsymbol{R}_{x,y}^{x',y'} \boldsymbol{v}_{x,y}^{x',y'} - \frac{1}{2} \beta {g_{x,y}}^2 - \frac{1}{2} \beta {g_{x',y'}}^2$$
$$- \frac{1}{2} \sum_{(x'',y'') \in \mathbf{c}_{x,y} \setminus (x',y')} \left( \mu_{x,y}^{x'',y''} \right)^2 \lambda_{x,y}^{x'',y''} - \frac{1}{2} \sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} \left( \mu_{x',y'}^{x'',y''} \right)^2 \lambda_{x',y'}^{x'',y''}, \tag{50}$$

respectively. By substituting equations (49) and (50) into equation (37), we obtain the approximate value of the free energy $\mathcal{F} = -\ln \mathcal{Z}$ in the Bethe approximation.

In dealing with equations (42) and (43), the free energy $\mathcal{F}$ is obtained by solving the simultaneous fixed-point equations (42) and (43) numerically and by substituting the set of solutions $\left\{ \lambda_{x,y}^{x',y'}, \mu_{x,y}^{x',y'} \middle| (x', y') \in \mathbf{c}_{x,y}, (x, y) \in \Omega \right\}$ into the right-hand side of equation (37) with equations (49) and (50). The simultaneous fixed-point equations (42) and (43) are solved by the following iterative algorithm.

Iterative algorithm for solving the simultaneous fixed-point equations (42) and (43)

*Step 1*. Set $r \leftarrow 0$ as an initial value.

*Step 2*. Update $r \leftarrow r + 1$ and

$$a_{x,y}^{x',y'}(r) \leftarrow \left( \frac{1}{\alpha} + \frac{1}{\beta + \sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} a_{x',y'}^{x'',y''}(r-1)} \right)^{-1}, \tag{51}$$

$$b_{x,y}^{x',y'}(r) \leftarrow \frac{\beta g_{x',y'} + \sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} b_{x',y'}^{x'',y''}(r-1) a_{x',y'}^{x'',y''}(r-1)}{\beta + \sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} a_{x',y'}^{x'',y''}(r-1)}. \tag{52}$$

*Step 3*. Update $R \leftarrow r$, $\lambda_{x,y}^{x',y'} \leftarrow a_{x,y}^{x',y'}(R)$ and $\mu_{x,y}^{x',y'} \leftarrow b_{x,y}^{x',y'}(R)((x, y) \in \Omega)$. Stop if, for pre-specified $\varepsilon$,

$$\sum_{(x,y) \in \Omega} \sum_{(x',y') \in \mathbf{c}_{x,y}} \left( \left| a_{x,y}^{x',y'}(r) - a_{x,y}^{x',y'}(r-1) \right| + \left| b_{x,y}^{x',y'}(r) - b_{x,y}^{x',y'}(r-1) \right| \right) < \varepsilon, \tag{53}$$

and go to step 2 otherwise.

Again, it is usually adequate to set $\varepsilon = 10^{-6}$. In the denominators of equations (51) and (52), the summations $\sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} a_{x',y'}^{x'',y''}(r-1)$ and $\sum_{(x'',y'') \in \mathbf{c}_{x',y'} \setminus (x,y)} b_{x',y'}^{x'',y''}(r-1) a_{x',y'}^{x'',y''}(r-1)$ can be evaluated in $\mathcal{O}(1)$ time per pair of pixels $(x, y)$ and $(x', y')$, because the number of elements in the set $\mathbf{c}_{x',y'} \setminus (x, y)$ is equal to 3 per pair of pixels. Hence the iterative algorithm for solving the simultaneous fixed-point equations (42) and (43) requires a total of $\mathcal{O}(|\Omega|)$ computations per update.

The results obtained for $\mathcal{F} = -\ln \mathcal{Z}$ by setting $\beta = 1/\sigma^2$ and $\beta = 0$ correspond to $\mathcal{F}_{\mathrm{POS}}(\boldsymbol{g}, \alpha, \sigma) \equiv -\ln \mathcal{Z}_{\mathrm{POS}}(\boldsymbol{g}, \alpha, \sigma)$ and $\mathcal{F}_{\mathrm{PR}}(\alpha) \equiv -\ln \mathcal{Z}_{\mathrm{PR}}(\alpha)$, respectively. By substituting these results in equation (11), we can calculate the approximate values of the marginal likelihood $\mathcal{P}(\boldsymbol{G} = \boldsymbol{g} | \alpha, \sigma)$ for any values of $\alpha$ and $\sigma$ in the mean-field and Bethe approximations.

## 4. Numerical experiments

In this section, we present the results of some numerical experiments for the restoration of grey-level images. The optimal values for the hyperparameters, $(\widehat{\alpha}, \widehat{\sigma})$, are determined by

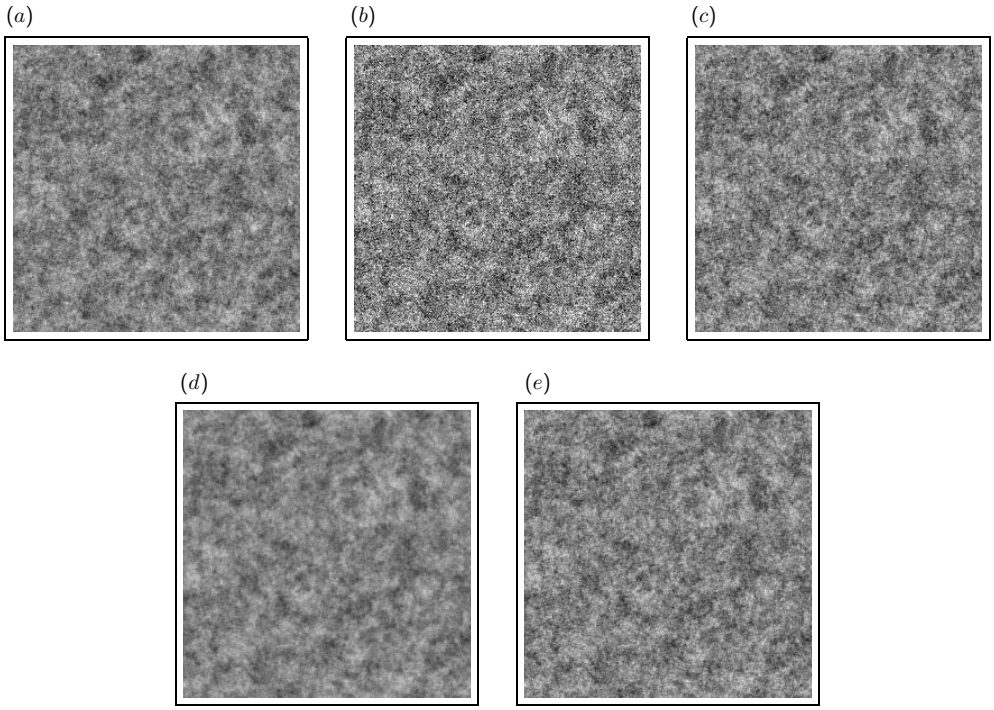


**Figure 1.** Image restoration based on the Gaussian graphical model. The original image $\boldsymbol{f}$ is generated from the *a priori* probability distribution (3). (*a*) Original image $\boldsymbol{f}$ ($\alpha = 0.0010$). (*b*) Degraded image $\boldsymbol{g}$ ($\sigma = 40$). (*c*) Restored image $\widehat{\boldsymbol{f}}$ obtained by the mean-field approximation ($\widehat{\alpha} = 0.000\,298, \widehat{\sigma} = 29.054$). (*d*) Restored image $\widehat{\boldsymbol{f}}$ obtained by the Bethe approximation ($\widehat{\alpha} = 0.000\,784, \widehat{\sigma} = 37.773$). (*e*) Restored image $\widehat{\boldsymbol{f}}$ obtained by the exact calculation ($\widehat{\alpha} = 0.001\,086, \widehat{\sigma} = 39.362$).

means of maximum marginal likelihood estimation, and the values of the marginal likelihood are calculated by using the exact expression (12), the mean-field approximation and the Bethe approximation.

In the present paper, our framework for Bayesian image restoration uses a Gaussian graphical model and considers a continuous random variable for the intensity at each pixel in the original and the degraded images. However, in practical images in computer vision, the intensity of light at each pixel is represented as an integer chosen from the set $\{0, 1, \ldots, 255\}$. In our numerical experiments, we apply the framework established in the previous sections to images consisting of integers $\{0, 1, \ldots, 255\}$. Instead of equation (10), we use

$$\widehat{f}_{x,y} \equiv \arg \min_{n=0,1,\ldots,255} \left( n - \int z_{x,y} \mathcal{P}(\boldsymbol{F} = \boldsymbol{z} | \boldsymbol{G} = \boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma}) \, \mathrm{d}\boldsymbol{z} \right)^2 . \tag{54}$$

To evaluate restoration performance quantitatively, 20 original images $\boldsymbol{f}$ are simulated from the *a priori* probability density (3) for the Gaussian graphical model. We produce three degraded images $\boldsymbol{g}$ from each original image $\boldsymbol{f}$ by means of the degradation process (2) for $\sigma = 30, 40$ and 50, respectively. By applying the exact expression in section 2 and iterative algorithms for obtaining the mean-field and Bethe approximations in section 4 to each degraded image $\boldsymbol{g}$, we obtain estimates of the hyperparameters $\widehat{\alpha}$ and $\widehat{\sigma}$ and the restored image $\widehat{\boldsymbol{f}}$ for each degraded image $\boldsymbol{g}$. For the case $\alpha = 0.0010$ and $\sigma = 40$, one of the numerical experiments is shown in figure 1. In order to illustrate the maximization of the marginal likelihood, we

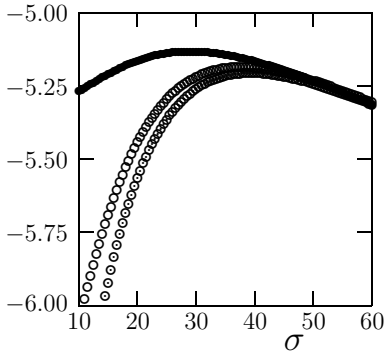


**Figure 2.** $\sigma$-dependence of the logarithm of the marginal likelihood per pixel, $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \sigma) \equiv \frac{1}{|\Omega|} \ln \mathcal{P}(\boldsymbol{G} = \boldsymbol{g}|\widehat{\alpha}, \sigma)$, for the image restorations shown in figure 1, where $\widehat{\alpha} = 0.000\,298$, $0.000\,784$, $0.001\,086$ for the mean-field approximation, the Bethe approximation and the exact calculation, respectively. The full circles, the open circles and the double circles correspond to the mean-field approximation, the Bethe approximation and the exact calculation, and give maximum values at $\sigma = 29.054$, 37.773, 39.362, respectively.

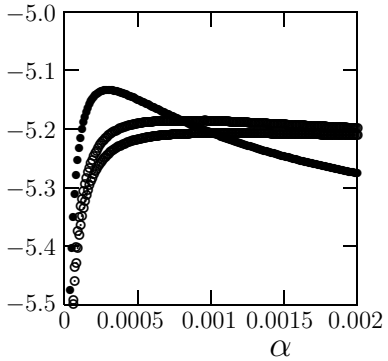


**Figure 3.** $\alpha$-dependence of the logarithm of the marginal likelihood per pixel, $\mathcal{L}(\boldsymbol{g}, \alpha, \widehat{\sigma}) \equiv \frac{1}{|\Omega|} \ln \mathcal{P}(\boldsymbol{G} = \boldsymbol{g}|\alpha, \widehat{\sigma})$, for the image restorations shown in figure 1, where $\widehat{\sigma} = 29.054$, 37.773, 39.362 for the mean-field approximation, the Bethe approximation and the exact calculation, respectively. The full circles, the open circles and the double circles correspond to the mean-field approximation, the Bethe approximation and the exact calculation, and give maximum values at $\alpha = 0.000\,298$, $0.000\,784$, $0.001\,086$, respectively.

display, in figures 2 and 3, $\sigma$- and $\alpha$-dependences of $\mathcal{L}(\boldsymbol{g}, \alpha, \sigma) \equiv \frac{1}{|\Omega|} \ln \mathcal{P}(\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma)$, corresponding to the image restorations shown in figure 1, respectively. In figure 2 $\alpha$ is fixed at $\hat{\alpha}$ and in figure 3 $\sigma$ is fixed at $\hat{\sigma}$. From the 20 degraded images $\boldsymbol{g}$ and the corresponding restored images $\widehat{\boldsymbol{f}}$, we calculate 95% confidence intervals for the hyperparameters, $\alpha$ and $\sigma$, and the values of the mean-squared error $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$, defined by

$$d(\boldsymbol{f}, \widehat{\boldsymbol{f}}) \equiv \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} (f_{x,y} - \widehat{f}_{x,y})^2. \tag{55}$$

These confidence intervals are given in table 1. Although clearly the true values of the hyperparameters, $\alpha$ and $\sigma$, lie outside the 95% confidence intervals based on the mean-field and the Bethe approximations, the confidence interval based on the Bethe approximation is considerably closer to the true value and to the confidence interval based on the exact calculation than that provided by the mean-field approximation. Thus there is a substantial


**Figure 4.** Original images: (*a*) Lena; (*b*) Mandrill.

**Table 1.** Approximate 95% confidence intervals for the hyperparameters, $\alpha$ and $\sigma$, and the values of $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma}) \equiv \frac{1}{|\Omega|} \ln \mathcal{P}(\boldsymbol{G} = \boldsymbol{g}|\widehat{\alpha}, \widehat{\sigma})$ and $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ obtained for some degraded images $\boldsymbol{g}$, which are produced for $\sigma = 30, 40, 50$ from 20 original images $\boldsymbol{f}$. The 20 original images are generated by Monte Carlo simulation from the *a priori* probability distribution (3) for the Gaussian graphical model ($\alpha = 0.001$). The hyperparameters are estimated by applying the mean-field approximation, the Bethe approximation and the exact method to maximum marginal likelihood estimation.

| | Mean-field approximation | Bethe approximation | Exact |
|---|---|---|---|
| | | $\sigma = 30$ | |
| $\widehat{\alpha}$ | [0.000 4100, 0.000 4133] | [0.000 8565, 0.000 8718] | [0.001 0117, 0.001 0372] |
| $\widehat{\sigma}$ | [21.595, 21.769] | [28.585, 28.769] | [29.58 360, 29.78 772] |
| $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma})$ | [−4.9118, −4.9085] | [−4.9724, −4.9694] | [−4.9955, −4.9925] |
| $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ | [367.32, 370.28] | [244.17, 245.18] | [242.25, 243.46] |
| | | $\sigma = 40$ | |
| $\widehat{\alpha}$ | [0.000 2941, 0.000 2962] | [0.000 7603, 0.000 7728] | [0.001 0435, 0.001 0689] |
| $\widehat{\sigma}$ | [29.019, 29.141] | [37.760, 37.857] | [39.353, 39.460] |
| $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma})$ | [−5.137 05, −5.135 11] | [−5.190 08, −5.188 11] | [−5.211 23, −5.209 23] |
| $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ | [536.42, 540.10] | [301.38, 302.26] | [295.58, 296.81] |
| | | $\sigma = 50$ | |
| $\widehat{\alpha}$ | [0.000 2219, 0.000 2207] | [0.000 6705, 0.000 6805] | [0.001 0755, 0.001 1074] |
| $\widehat{\sigma}$ | [36.200, 36.317] | [46.551, 46.646] | [48.782, 48.669] |
| $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma})$ | [−5.320 64, −5.319 04] | [−5.368 48, −5.366 89] | [−5.387 80, −5.386 18] |
| $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ | [722.26, 726.68] | [347.63, 349.53] | [337.65, 335.41] |

improvement in accuracy if we use the Bethe approximation rather than the mean-field approximation for hyperparameter estimation.

We then performed numerical experiments based on the artificial images in figure 4. Degraded images $\boldsymbol{g}$ are produced from the original images $\boldsymbol{f}$ with $\sigma = 30, 40$ and 50. The degraded images $\boldsymbol{g}$ for $\sigma = 40$ are shown in figure 5. The image restorations created by means of the mean-field approximation, the Bethe approximation and the exact calculation for the Gaussian graphical model are shown in figures 6(*a*)–(*c*) and figures 7(*a*)–(*c*). We give in table 2 the estimates, $\widehat{\sigma}$ and $\widehat{\alpha}$, of the hyperparameters, and the values of the mean-squared error $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$. For the practical images, we again find that the estimates of the hyperparameters, $\widehat{\alpha}$ and $\widehat{\sigma}$, provided by the Bethe approximation are closer to the exact results than those based on the mean-field approximation. Particularly in the context of the mean-squared error measure, the Bethe approximation again provides a clear improvement over the mean-field approximation and does nearly as well as the exact method. Of course we have to bear in

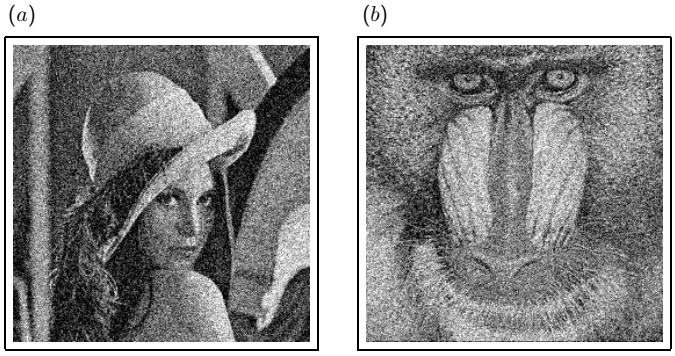


**Figure 5.** Degraded images ($\sigma = 40$): (*a*) Lena; (*b*) Mandrill.

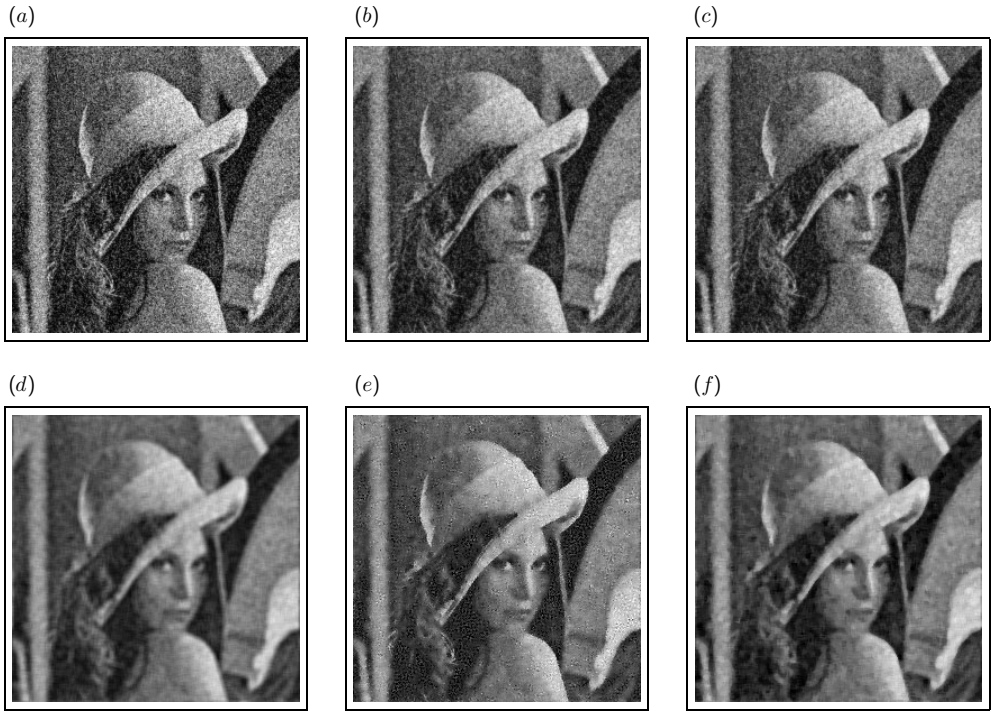


**Figure 6.** Restored images $\widehat{f}$ obtained from the degraded image $g$ given in figure 5(*a*). (*a*) Mean-field approximation. (*b*) Bethe approximation. (*c*) Exact. (*d*) Lowpass filter with window size $5 \times 5$. (*e*) Wiener filter with window size $5 \times 5$. (*f*) Median filter with window size $5 \times 5$.

mind that real images are not generated from the model for which the above exact method is correct, but the performance of the Bethe approximation is nevertheless very encouraging.

Another important point is to note that the probabilistic methods appear to do better than conventional filter methods in image processing, as we now illustrate. For the degraded images shown in figure 5, we apply the $(2n + 1)\times(2n + 1)$ lowpass filter, the $(2n + 1)\times(2n + 1)$ Wiener filter and the $(2n + 1)\times(2n + 1)$ median filter for $n = 1, 2$ [25, 26]. In the $(2n + 1)\times(2n + 1)$ lowpass filter, the output $\widehat{f}_{x,y}$ is the average of the inputs $\{g_{x',y'} | x - n \leqslant x' \leqslant x + n,\ y - n \leqslant y' \leqslant y + n\}$. In the $(2n+1)\times(2n+1)$ median filter,

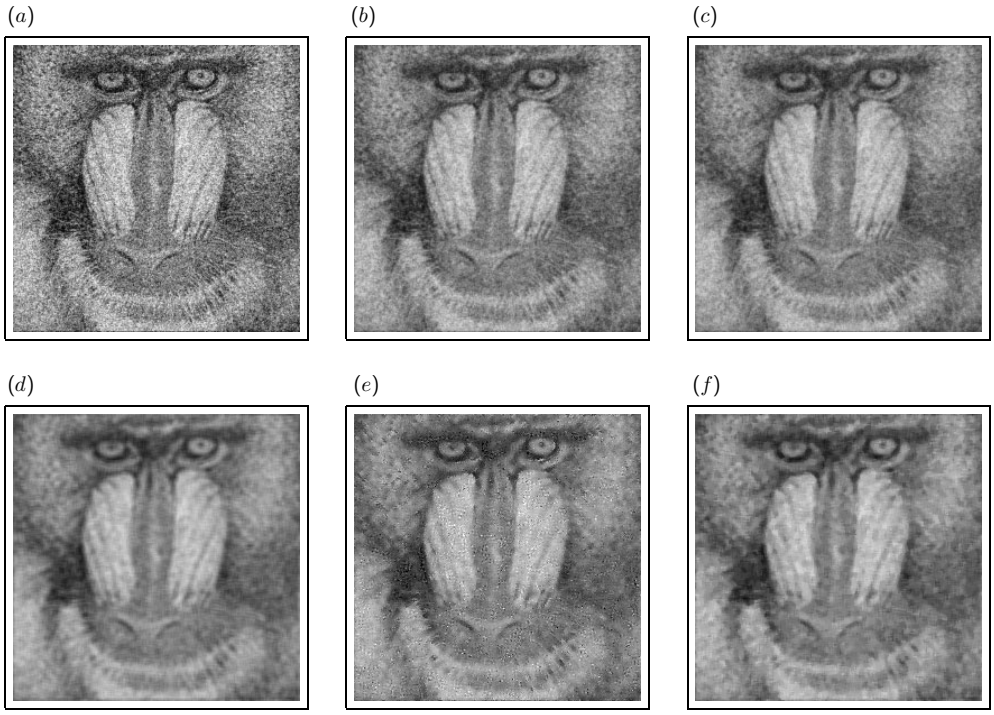


**Figure 7.** Restored images $\widehat{\boldsymbol{f}}$ obtained from the degraded image $\boldsymbol{g}$ given in figure 5(*b*). (*a*) Mean-field approximation. (*b*) Bethe approximation. (*c*) Exact. (*d*) Lowpass filter with window size 5×5. (*e*) Wiener filter with window size 5 × 5. (*f*) Median filter with window size 5 × 5.

the output $\widehat{f}_{x,y}$ is the median of the inputs $\{g_{x',y'}|x-n \leqslant x' \leqslant x+n, y-n \leqslant y' \leqslant y+n\}$. In the $(2n+1)\times(2n+1)$ Wiener filter, the output $\widehat{f}_{x,y}$ is

$$\widehat{f}_{x,y} = a_{x,y} + \frac{b_{x,y} - \nu}{b_{x,y}}(g_{x,y} - a_{x,y}), \tag{56}$$

where $a_{x,y}$ and $b_{x,y}$ are the 'sample' average and the 'sample' variance of the inputs $\{g_{x',y'}|x-n \leqslant x' \leqslant x+n, y-n \leqslant y' \leqslant y+n\}$, respectively, and $\nu$ is the average of $\{b_{x,y}|(x, y) \in \Omega\}$. Some of the restored images $\widehat{\boldsymbol{f}}$ obtained by applying the lowpass filter, the Wiener filter and the median filter to the degraded images given in figure 5 are shown in figures 6(*d*)–(*g*) and figures 7(*d*)–(*g*). The mean-squared errors $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ for the restored images obtained by means of the conventional filters are given in table 3. The (3 × 3) filters cannot erase the noise to a satisfactory level in terms of mean-squared error. If (5 × 5) filters are used, the noise can be erased adequately, but the restored image is blurred. It is obvious that, in 'Mandrill', the results obtained by the approach in this paper, based on a prior Gaussian graphical model, are better than those based on conventional filters. However, the situation with 'Lena' seems to be different. Comparison of the results for 'Lena' in tables 2 and 3 shows that the mean-squared errors $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ of the mean-field approximation, the Bethe approximation and the exact calculation are all greater than that of the (5 × 5) lowpass filter.

The Tikhonov method is also often applied to grey-level image restoration in conventional image processing [27]. The algorithm is based on the framework of the constrained least mean square filter [28]. For a given degraded image $\boldsymbol{g} = \{g_{x,y}|(x, y) \in \Omega\}$ in which the degradation

**Table 2.** Estimates of hyperparameters, $\widehat{\alpha}$ and $\widehat{\sigma}$, and the values of $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma}) \equiv \frac{1}{|\Omega|} \ln \mathcal{P}(\boldsymbol{G} = \boldsymbol{g} | \widehat{\alpha}, \widehat{\sigma})$ and $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ obtained by using the mean-field approximation (MFA), the Bethe approximation and the exact calculation. The degraded images $\boldsymbol{g}$ are produced for $\sigma = 30, 40, 50$ from the original images $\boldsymbol{f}$ in figure 4.

| $\sigma$ | $d(\boldsymbol{f}, \boldsymbol{g})$ | Method | $\widehat{\alpha}$ | $\widehat{\sigma}$ | $\mathcal{L}(\boldsymbol{g}, \widehat{\alpha}, \widehat{\sigma})$ | $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ |
|---|---|---|---|---|---|---|
| | | | Lena | | | |
| | | MFA | 0.000 393 | 16.762 | −4.829 57 | 433 |
| 30 | 813 | Bethe | 0.000 546 | 21.716 | −4.903 78 | 279 |
| | | Exact | 0.000 558 | 21.976 | −4.928 85 | 272 |
| | | MFA | 0.000 288 | 24.417 | −5.063 65 | 593 |
| 40 | 1409 | Bethe | 0.000 484 | 31.315 | −5.128 98 | 324 |
| | | Exact | 0.000 517 | 31.960 | −5.153 67 | 306 |
| | | MFA | 0.000 223 | 31.291 | −5.243 89 | 765 |
| 50 | 2109 | Bethe | 0.000 437 | 39.826 | −5.302 62 | 375 |
| | | Exact | 0.000 493 | 40.878 | −5.326 48 | 346 |
| | | | Mandrill | | | |
| | | MFA | 0.000 358 | 19.518 | −4.911 11 | 425 |
| 30 | 861 | Bethe | 0.000 645 | 26.846 | −4.979 30 | 262 |
| | | Exact | 0.000 708 | 27.636 | −5.003 78 | 255 |
| | | MFA | 0.000 270 | 27.453 | −5.132 21 | 591 |
| 40 | 1512 | Bethe | 0.000 613 | 36.469 | −5.190 71 | 325 |
| | | Exact | 0.000 757 | 37.857 | −5.213 51 | 315 |
| | | MFA | 0.000 211 | 34.771 | −5.309 52 | 766 |
| 50 | 2291 | Bethe | 0.000 573 | 45.276 | −5.361 27 | 380 |
| | | Exact | 0.000 802 | 47.156 | −5.382 26 | 364 |

**Table 3.** Values of $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ obtained by using the lowpass filters, Wiener filters and median filters with window sizes $3 \times 3$ and $5 \times 5$. The degraded images $\boldsymbol{g}$ are produced for $\sigma = 30, 40, 50$ from the original images $\boldsymbol{f}$ in figure 4.

| $\sigma$ | $d(\boldsymbol{f}, \boldsymbol{g})$ | Window size | Lowpass | Wiener | Median |
|---|---|---|---|---|---|
| | | | Lena | | |
| 30 | 813 | $3 \times 3$ | 185 | 492 | 220 |
| | | $5 \times 5$ | 232 | 369 | 202 |
| 40 | 1409 | $3 \times 3$ | 258 | 492 | 339 |
| | | $5 \times 5$ | 268 | 369 | 259 |
| 50 | 2109 | $3 \times 3$ | 350 | 903 | 484 |
| | | $5 \times 5$ | 311 | 518 | 320 |
| | | | Mandrill | | |
| 30 | 861 | $3 \times 3$ | 312 | 653 | 370 |
| | | $5 \times 5$ | 384 | 485 | 401 |
| 40 | 1512 | $3 \times 3$ | 447 | 848 | 488 |
| | | $5 \times 5$ | 411 | 545 | 447 |
| 50 | 2291 | $3 \times 3$ | 478 | 1080 | 439 |
| | | $5 \times 5$ | 447 | 627 | 504 |

takes the form of additive white Gaussian noise with mean 0 and variance $\sigma^2$, the most basic constrained least mean square filter is defined by

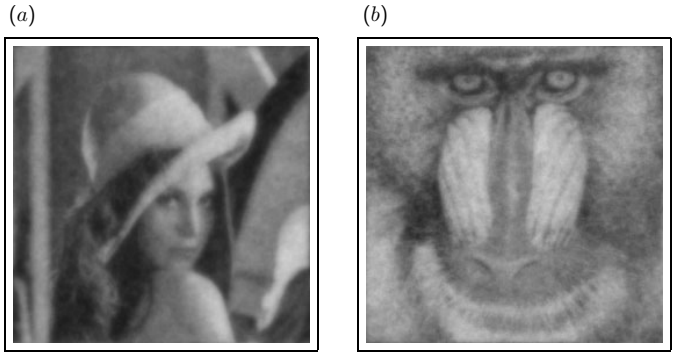


**Figure 8.** Restored images $\widehat{\boldsymbol{f}}$ obtained from the degraded image $\boldsymbol{g}$ by using the constrained least mean square filter. (*a*) Lena. (*b*) Mandrill.

**Table 4.** Values of $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ obtained by using the constrained least mean square filter. The degraded images $\boldsymbol{g}$ are produced for $\sigma = 30, 40, 50$ from the original images $\boldsymbol{f}$ in figure 4.

| | $\sigma = 30$ | $\sigma = 40$ | $\sigma = 50$ |
|---|---|---|---|
| Lena | 273 | 414 | 651 |
| Mandrill | 297 | 372 | 458 |

$$\widehat{\boldsymbol{f}} = \arg \min_{\boldsymbol{z}: d(\boldsymbol{z},\boldsymbol{g})=\sigma^2} \left( \sum_{(x,y)\in\Omega} ((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2) \right), \tag{57}$$

where $d(\boldsymbol{z}, \boldsymbol{g})$ is defined by equation (55). Some of the restored images $\widehat{\boldsymbol{f}}$ obtained by applying the constrained least mean square filter to the degraded images in figure 5 are shown in figure 8. The mean-squared errors $d(\boldsymbol{f}, \widehat{\boldsymbol{f}})$ for the restored images obtained by means of the constrained least mean square filter are given in table 4.

Before closing the present section, we report the number of iterations, $R$, and the computation times necessary for the iterative algorithms for solving the simultaneous fixed-point equations in the mean-field and Bethe approximations. We performed the present numerical experiments on a SONY VAIO PCG-GRT77E/P personal computer with a Pentium4 Processor and 1 GByte of memory. For example, for solving the simultaneous fixed-point equations, given $(\widehat{\alpha}, \widehat{\sigma})$, in the production of the results for figures 7(*a*) and (*b*) with the mean-field and Bethe approximations, the numbers of iterations are $R = 15$ and $R = 42$, respectively, and the corresponding computational times are 0.688 s and 10.391 s, respectively. For the lowpass, median and the Wiener filters with window sizes $5 \times 5$, the computation times are approximately 0.1 s, 0.3 s and 0.1 s, respectively, and the computation time for obtaining each result for the constrained least mean square filter is approximately 3 s. We expect that the computation times for the mean-field and Bethe approximations can be significantly reduced by improving the algorithms, especially the procedure for calculating $\widehat{\alpha}$ and $\widehat{\sigma}$, which is quite time-consuming.

## 5. Concluding remarks

In this paper, we have studied the accuracy of hyperparameter estimation based on the Bethe approximation. The estimates of the hyperparameters are determined so as to maximize the marginal likelihood. The marginal likelihood can be expressed in terms of the free energies

of the prior probabilistic model and the posterior probabilistic model. Hence the marginal likelihood can be calculated approximately by using the Bethe approximation or the mean-field approximation. The main purpose of this paper is to examine the accuracy of the hyperparameter estimates as obtained by maximizing the marginal likelihood as calculated approximately by means of the Bethe approximation. In the calculation, we have adopted a Gaussian graphical model, representing local spatial correlation, as the prior probabilistic model in Bayes' formula. This provides a probabilistic model for image restoration which can be treated analytically by using the multi-dimensional Gauss integral formula and the discrete Fourier transformation. We have therefore been able to compare the results based on the mean-field and Bethe approximations with those from the exact calculation. In the estimation of the hyperparameters, biases, relative to the exact results, are incurred with the Bethe approximation and the mean-field approximation. These tendencies appear not only in the context of practical real-world images but also with images generated from the proposed *a priori* probability distribution (3). However, the results obtained by the Bethe approximation represent substantial improvements over those obtained by the mean-field approximation.

As we have already mentioned in section 1, Weiss and Freeman [7] have pointed out that, in the Bethe approximation for the Gaussian graphical model, the accuracy of the average is good but that of the variance is not adequate. We believe that this lack of accuracy for the variance leads to the biases in the estimates of the hyperparameters, compared to the exact results, in the experiments.

Other models exist which are applicable to probabilistic image processing and which can be treated analytically by using the multi-dimensional Gauss formula and the discrete Fourier transformation [17–19]. It would be interesting to investigate how well the Bethe approximation and the mean-field approximation perform, relative to the exact methods, in such solvable models. This is a problem for the future.

Though each pixel intensity takes a real value in the range $(-\infty, +\infty)$ in the present paper, we can also consider the case in which the light intensity $f_{x,y}$ is not real valued but takes an integer value from the set $\{0, 1, 2, \ldots, Q-1\}$ for a fixed positive and finite integer $Q$. In this case, it is hard to calculate the marginal likelihood $\mathcal{P}(\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma)$, and Tanaka *et al* [29, 30] investigated the maximization of the marginal likelihood with the Bethe approximation in this context. Moreover, Tanaka and Titterington [31] compared the results obtained by the Bethe approximation with those obtained by the mean-field approximation, and showed that the Bethe approximation substantially improved upon the mean-field approximation. In [29–31], the simultaneous fixed-point equations for $\left\{\mathcal{M}_{x,y}^{x',y'}(\xi)\middle|(x, y) \in \Omega, (x', y') \in \mathbf{c}_{x,y}\right\}$ in the Bethe approximation for any finite positive integer $Q$ have been shown to take the form

$$\mathcal{M}_{x,y}^{x',y'}(\xi) = \frac{\sum_{\xi'=0}^{Q-1}\phi(\xi, \xi')\psi_{x',y'}(\xi')\prod_{(x'',y'')\in \boldsymbol{c}_{x',y'}\setminus(x,y)}\mathcal{M}_{x',y'}^{x'',y''}(\xi')}{\sum_{\xi'=0}^{Q-1}\sum_{\xi''=0}^{Q-1}\phi(\xi'', \xi')\psi_{x',y'}(\xi')\prod_{(x'',y'')\in \boldsymbol{c}_{x',y'}\setminus(x,y)}\mathcal{M}_{x',y'}^{x'',y''}(\xi')} \tag{58}$$

instead of equation (36). In this case, the simultaneous fixed-point equations (58) for each pair of pixels $(x, y)$ and $(x', y')$ can be evaluated in $\mathcal{O}(Q^2)$ computations, which implies that a total of $\mathcal{O}(Q^2|\Omega|)$ computations are required per update. In [29–31], we treated the case of $Q = 4$ at most. If we consider the case of $Q = 256$, we need unrealistic computation time and large amounts of memory and this is still beyond the capacity of present computers. In the present paper, intensities with 256 grey levels are regarded as continuous variables and we treat practical digital images by means of the Gaussian graphical models in equations (3) and (5). It is a future problem to compare the present results with those from the Bethe approximation for $Q = 256$.

In the Bethe approximation, the approximate free energy for every probabilistic model in equation (16) is given by equation (37). The simultaneous fixed-point equations (34)–(36) are equivalent to the extremum conditions of the Bethe free energy $\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1} \middle| (x, y) \in \Omega\right\}\right]$ with respect to the marginal probability distributions $\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1} \middle| (x, y) \in \Omega\right\}$ under the constraints (29) and (30). However, it is known that the Bethe free energy $\mathcal{F}_{\mathrm{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1} \middle| (x, y) \in \Omega\right\}\right]$ does not provide a guaranteed bound for the true free energy $\mathcal{F} = -\ln\mathcal{Z}$, whereas a mean-field free energy does provide such a bound [2]. Furthermore, in some cases the solution of the simultaneous fixed-point equations (34)–(36) corresponds not to a local minimum but to a saddle point of the Bethe free energy [41]. In spite of that, the present scheme provides satisfactory results. The Bethe approximation often gives us poor results for some Ising models with frustration effects [32, 33]. Frustration effects in probabilistic models cause poor results in the Bethe approximation. The present *a priori* and *a posteriori* probabilistic models (3) and (5) have no frustration effects since they correspond to the Gaussian graphical ones with spatially uniform ferromagnetic interactions. This seems to be a qualitative explanation of why satisfactory results are obtained in the present scheme in spite of the problems that can arise with the Bethe approximation.

One extension of the Bethe approximation is the cluster variation method [34, 35]. Recently, some researchers in statistical mechanics and computer science have applied the cluster variation method to the probabilistic information sciences [9, 36]. Tanaka, Inoue and Titterington have set out a general scheme for the maximization of marginal likelihood based on the cluster variation method [30]. However, the scheme has only been applied to simple binary cases and not yet to practical computer vision contexts. On the other hand, some attempts to apply new advanced mean-field methods to probabilistic information processing problems have combined the Bethe approximation with the linear response formula or have used the Thouless–Anderson–Palmer (TAP) equation [37, 38]. More detailed investigation of the application of these methods to probabilistic image processing is also the subject of future work.

The Bethe approximation has already been applied also to error-correcting codes in coding theory [8] and to the code division multiple access (CDMA) systems in wireless communication [39]. Moreover, Kanter and Kfir proposed a statistical-mechanical joint source–channel decoder based on the transfer matrix method [40]. This idea is also applicable to Bayesian image restoration. It may be useful to compare the present method with the application of a statistical-mechanical joint source–channel decoder. This is left as a future problem.

The main limitation of the present approach is that the degraded images $\boldsymbol{g}$ are obtained by simply adding white Gaussian noise to the original image, without considering modifications to represent the optical device used for image acquisition. If the impulse response of the instrument is considered, the conditional probability density $\mathcal{P}(\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f})$ describing the degradation process would assume a more complicated form. In such a case, we have to treat equation (58), and not equations (42) and (43), as the relevant simultaneous fixed-point equations. Though it is possible to solve such simultaneous fixed-point equations numerically, a total of $\mathcal{O}(Q^2|\Omega|)$ operations are required per update in the iterative algorithm. The whole computation would become much more involved and we would have to reduce the computational complexity by combining certain other approximations with the Bethe approximation. This also is a future problem.

As a more practical image restoration scheme, we can introduce into the prior probabilistic model a line field $\boldsymbol{u} = \left\{u_{x,y}^{x+1,y}, u_{x,y}^{x+1,y} \middle| (x, y) \in \Omega\right\}$, so that instead of equation (3) we have

$$\mathcal{P}(\boldsymbol{F}=\boldsymbol{f}|\alpha,\gamma)\equiv\frac{1}{\mathcal{Z}_{\mathrm{PR}}(\alpha,\gamma)}\sum_{\boldsymbol{u}}\prod_{(x,y)\in\Omega}\exp\left(-\frac{1}{2}\alpha\bigl(1-u_{x,y}^{x+1,y}\bigr)\bigl((f_{x,y}-f_{x+1,y})^2-\gamma^2\bigr)\right.$$
$$\left.-\frac{1}{2}\alpha\bigl(1-u_{x,y}^{x,y+1}\bigr)\bigl((f_{x,y}-f_{x,y+1})^2-\gamma^2\bigr)\right). \tag{59}$$

Here a random variable $u_{x,y}^{x',y'}$ representing the edge is assigned to each nearest-neighbour pair of pixels, $(x, y)$ and $(x', y')$. The states $u_{x,y}^{x',y'}=1$ and $u_{x,y}^{x',y'}=0$ correspond to the existence or non-existence of an edge, respectively. The summation $\sum_{\boldsymbol{u}}$ ranges over all possible values of the line fields $u_{x,y}^{x',y'}$ of all nearest-neighbour pairs of pixels, $(x, y)$ and $(x', y')$. In this case, it is hard to calculate the marginal likelihood, $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\gamma,\sigma)\equiv\int\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{z},\sigma)\mathcal{P}(\boldsymbol{F}=\boldsymbol{z}|\alpha,\gamma)\,\mathrm{d}\boldsymbol{z}$, analytically. For the posterior probabilistic model given by $\mathcal{P}(\boldsymbol{F}=\boldsymbol{f}|\boldsymbol{G}=\boldsymbol{g},\alpha,\gamma,\sigma)\equiv\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{f},\sigma)\mathcal{P}(\boldsymbol{F}=\boldsymbol{f}|\alpha,\gamma)/\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\gamma,\sigma)$, one of the present authors has developed the belief propagation algorithm based on the cluster variation method [42]. However, no one has yet succeeded in maximizing the marginal likelihood $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\gamma,\sigma)$ with respect to $\alpha$, $\gamma$ and $\sigma$ based on any cluster variation method, including the Bethe approximation. Moreover, we may have to consider models including interactions within the line field $\boldsymbol{u}$ [13, 43]. Such a model is called a *Gauss–Markov random-field model*. Further application of the Bethe approximation and the cluster variation method (i.e. generalized belief propagation) to the maximization of the marginal likelihood $\mathcal{P}(\boldsymbol{G}=\boldsymbol{g}|\alpha,\gamma,\sigma)$ in the Gauss–Markov random-field model will be another subject of future research.

## Acknowledgments

The authors are grateful to Professor T Horiguchi of the Graduate School of Information Science, Tohoku University, for valuable discussions, and to the referees for their helpful comments. This work was partly supported by the Grants-In-Aid (no 14084203 and no 14084212) for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

[1] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing—An Introduction* (Oxford: Oxford University Press)
[2] Opper M and Saad D (ed) 2001 *Advanced Mean Field Methods—Theory and Practice* (Cambridge, MA: MIT Press)
[3] Pearl J 1988 *Probabilistic Reasoning in Intelligent Systems Networks of Plausible Inference* (San Mateo, CA: Morgan Kaufmann Publisher)
[4] Lauritzen S L 1995 *Graphical Models (Oxford Science Publication)* (Oxford: Oxford University Press)
[5] Morita T 1976 *Physica* A **83** 411
[6] Weiss Y 2000 *Neural Comput.* **12** 1
[7] Weiss Y and Freeman W T 2001 *Neural Comput.* **13** 2173
[8] Kabashima Y and Saad D 1998 *Europhys. Lett.* **44** 668
[9] Yedidia J S, Freeman W T and Weiss Y 2001 *Advances in Neural Information Processing Systems* **13** 689 (Cambridge, MA: MIT Press)
[10] Freeman W T, Jones T R and Pasztor E C 2002 *IEEE Comput. Graph. Appl.* **22** 56
[11] Derin H, Elliott H, Cristi R and Geman D 1984 *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 707
[12] Geman S and Geman D 1984 *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 721
[13] Chellappa R and Jain A (ed) 1993 *Markov Random Fields: Theory and Applications* (New York: Academic)
[14] Li S Z 1995 *Markov Random Field Modeling in Computer Vision* (Tokyo: Springer)
[15] Tanaka K 2002 *J. Phys. A: Math. Gen.* **35** R81

[16] Molina R 1994 *IEEE Trans. Pattern Anal. Mach. Intell.* **16** 1122
[17] Tsuzurugi J and Okada M 2002 *Phys. Rev.* E **66** 066704
[18] Tanaka K and Inoue J 2002 *IEICE Trans. Inf. Syst.* D **85** 546
[19] Tanaka K and Horiguchi T 2002 *Phys. Rev.* E **65** 046142
[20] Nishimori H 2000 *BUSSEI KENKYU* **73** 850
[21] MacKay D J 1992 *Neural Comput.* **4** 415
[22] Pryce J M and Bruce A D 1995 *J. Phys. A: Math. Gen.* **28** 511
[23] Marroquin J, Mitter S and Poggio T 1987 *J. Am. Stat. Assoc.* **82** 76
[24] Nishimori H and Wong K Y M 1999 *Phys. Rev.* E **60** 132
[25] Lim L S 1990 *Two-Dimensional Signal and Image Processing* (Englewood Cliffs, NJ: Prentice Hall)
[26] Gonzales R C and Woods R E 1992 *Digital Image Processing* (Reading, MA: Addison-Wesley)
[27] Hansen P C 1998 *Rank-Deficient and Discrete Ill-Posed Problems* (Philadelphia: SIAM)
[28] Hunt B R 1973 *IEEE Trans. Comput.* **22** 805
[29] Tanaka K, Inoue J and Titterington D M 2003 *J. Phys. A: Math. Gen.* **36** 11023
[30] Tanaka K, Inoue J and Titterington D M 2003 Neural networks for signal processing XIII *Proc. 2003 IEEE Signal Processing Society Workshop* vol 329
[31] Tanaka K and Titterington D M 2004 *Proc. 17th Int. Conf. on Pattern Recognition (Cambridge)* at press
[32] Morita T 1987 *Physica* A **141** 335
[33] Morita T 1988 *Phys. Lett.* A **132** 1
[34] Kikuchi R 1951 *Phys. Rev.* **81** 988
[35] Morita T 1972 *J. Math. Phys.* **13** 115
[36] Tanaka K 2003 *IEICE Tran. Inf. Syst.* D **86** 1228
[37] Welling M and Teh Y W 2003 *Artif. Intell.* **143** 19
[38] Welling M and Teh Y W 2004 *Neural Comput.* **16** 197
[39] Kabashima Y 2003 *J. Phys. A: Math. Gen.* **36** 11111
[40] Kanter I and Kfir H 2003 *Europhys. Lett.* **63** 310
[41] Heskes T 2003 *Adv. Neural Inf. Process. Syst.* **15** 359
[42] Tanaka K 2001 *Trans. Japan. Soc. Artif. Intell.* **16** 259
[43] Jeng F C and Woods J W 1991 *IEEE Trans. Signal Process.* **39** 683