Arnab Das and Bikas K. Chakrabarti (Eds.)

Quantum Annealing and Related Optimization Methods

– Monograph –

December 1, 2005

Springer
Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

Contents

1 Quantum Spin Glasses, Quantum Annealing and Probabilistic Information Processing

Jun-ichi Inoue

Graduate School of Information Science and Technology, Hokkaido University

Summary. We present several applications of quantum spin glasses (random field Ising model, Sherrington-Kirkpatrick model, Ising spin glasses with p-body interaction in a transverse field) to probabilistic information processing, especially to the problems of image restoration and error-correcting codes. As a related optimization method, quantum annealing is also introduced to these research fields and its performance is investigated by using the quantum Markov chain Monte Carlo method. After a short review of the previous work [J. Inoue, *Physical Review E* **63**, 046114 (2001)], we evaluate the performance of both the Maximum A Priori (MAP for short) and the Maximizer of Posterior Marginal (MPM for short) image restorations which are purely driven by quantum fluctuation (without any thermal fluctuation). The Nishimori-Wong condition, on which the best possible performance of the quantum MPM estimation is achieved, is derived as a condition on the effective amplitude of the transverse field. We show the lowest values of the bit-error rate for both the thermal and the quantum MPM estimations are exactly same. We next discuss an extension of the Sourlas codes by means of Ising spin glasses with p-body interactions in a transverse field. We investigate the tolerance of error-less (or quite low-error) ferromagnetic state to quantum uncertainties in the prior distribution. We find that there exist some critical amplitudes of the transverse field, and at the critical point the system changes from the low-error state to the poor error-correction state as the second order ($p = 2$) and as the first order ($p \geq 3$) phase transitions. The relation between the amplitude of the transverse field and the Shannon's information bound is also discussed in the limit of $p \to \infty$ for a given effective amplitude of the transverse field. We show that in this limit the Shannon's bound is not violated by the quantum fluctuation in the prior. In last part of this article, we apply quantum annealing, which is an optimization method based on quantum fluctuations, to the problem of image restoration. We compare the results of the thermal MAP and the quantum MAP estimations by using the simulated (thermal) and the quantum annealings, respectively. We find that a fine restoration of image is achieved by the quantum annealing and its performance measured by the bit-error rate is slightly superior to that of the simulated annealing.

1.1 Introduction

Recently, problems of information processing were investigated from statistical mechanical point of view [1]. Among them, image restoration (see [2, 3, 4] and references there in) and error-correcting codes [5] are most suitable subjects. In the field of the error-correcting codes, Sourlas [5] showed that the convolution codes can be constructed by infinite range spin-glasses Hamiltonian and the decoded message should correspond to the zero temperature spin configuration of the Hamiltonian. Ruján [6] suggested that the error of each bit can be suppressed if one uses finite temperature equilibrium states (sign of the local magnetization) as the decoding result, what we call the *MPM* (*Maximizer of Posterior Marginal*) estimate, instead of zero temperature spin configurations, and this optimality of the retrieval quality at a specific decoding temperature (this temperature is well known as the *Nishimori temperature* in the field of spin glasses) is proved by Nishimori [7].

The next remarkable progress in this direction was done by Nishimori and Wong [8]. They succeeded in giving a new procedure in order to compare the performance of the zero temperature decoding (statisticians call this strategy the *MAP* (*Maximum A Posteriori*) estimation) with that of the finite temperature decoding, the MPM estimation. They introduced an infinite range model of spin-glasses like the Sherrington-Kirkpatrick (SK) model [9] as an exactly solvable example. Kabashima and Saad [10] succeeded in constructing more practical codes, namely, low density parity check (LDPC) codes by using the spin glass model with finite connectivities. In these decoding process, one of the most important problems is how one obtains the minimum energy states of the effective Hamiltonian as quickly as possible. Geman and Geman [11] used simulated annealing [12] in the context of image restoration to obtain good recovering of the original image from its corrupted version. Recently, Tanaka and Horiguchi [13, 2] introduced a quantum fluctuation, instead of the thermal one, into the mean-field annealing algorithm and showed that performance of the image recovery is improved by controlling the quantum fluctuation appropriately during its annealing process. The attempt to use the quantum fluctuation to search the lowest energy states in the context of annealings by Markov chain Monte Carlo methods, what we call *quantum annealing*, is originally introduced by [14, 15] and its application to the combinatorial optimization problems including the ground state searching for several spin glass models was done by Kadowaki and Nishimori [16] and Santoro *et al* [17]. However, these results are restricted to research aided by computer simulations, although there exist some extensive studies on the Landou-Zener's model for the single spin problems [18, 19, 20].

Recently, the averaged case performance of the both MPM and MAP estimations for image restoration with quantum fluctuation was investigated by the present author [21] for the mean-field model. He also carried out the quantum Monte Carlo method to evaluate the performance for two dimensional pictures and found that the quantum fluctuation suppress the error

due to failing to set the hyperparameters effectively, however, the best possible value of the bit-error rate does not increases by the quantum fluctuation. In this result the quantum and the thermal fluctuations are combined in the MPM estimation (the effective temperature is unity). Therefore, it is important for us to revisit this problem and investigate to what extent the MPM estimation, which is based on pure quantum fluctuation and without any thermal one, works effectively.

In this article, we make this point clear and show that the best possible performance obtained by the MPM estimation, which is purely induced by quantum fluctuations, is exactly same as the results by the thermal MPM estimation. The Nishimori-Wong condition [7, 8] for the quantum fluctuation, on which the best possible performance is achieved, is also discussed. Moreover, we extend the Sourlas codes [5] by means of the spin glass model with p-spin interaction in a transverse field [22, 23] and discuss the tolerance of error-less (or quite low-error) state to the quantum uncertainties in the prior distribution. In last part of this article, we check the performance of the MAP and MPM image restorations predicted by the analysis of the mean-field infinite range model by using the quantum Markov chain Monte Carlo method [24] and the quantum annealing [14, 15, 16, 17].

This article is organized as follows. In the next section 2 and following section 3, we introduce our model system for image restoration and error-correcting codes. We also explain the relation between Bayesian inference and statistical mechanics. In section 4, we investigate the performance of the MAP and MPM estimations for these two problems by using the analysis of the infinite range model. In section 5, we carry out the quantum Markov chain Monte Carlo method and the quantum annealing to check the results we obtained from the analysis of the infinite range models. The final section is summary.

1.2 Bayesian statistics and information processing

In the field of signal processing or information science, we need to estimate the original message which is sent via email or fax. Usually, these massages are degraded by some noise and we should retrieve the original messages, and if possible, we send these messages not only as sequence of information bits but as some redundant information like *parity check*. In such problems, noise channels or statistical properties of the original message are specified by some appropriate probabilistic models. In this section, we explain the general definitions of our problems and how these problems link to statistical physics.

1.2.1 General definition of the model system

Let us suppose that the original information is represented by a configuration of Ising spins $\{\xi\} \equiv (\xi_1, \xi_2, \cdots, \xi_N)$ $(\xi_i = \pm 1, i = 1, \cdots, N)$ with probability

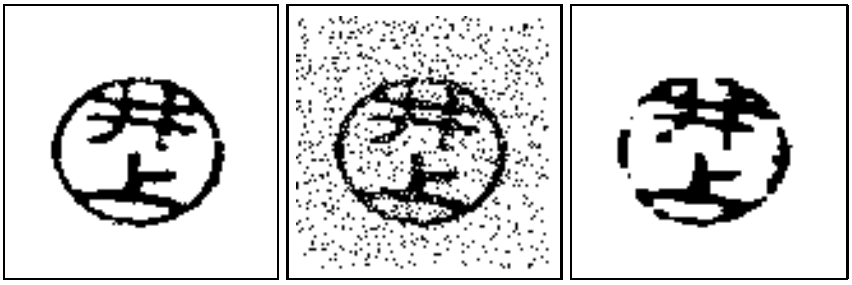

Fig. 1.1. A typical example of image data retrieval. From the left to the right, the original $\{\xi\}$, the degraded $\{\tau\}$ and the recovering $\{\sigma\}$ images. The above restored image was obtained by quantum annealing. The detailed account of this method will be explained and discussed in last part of this article.

$P(\{\xi\})$. Of course, if each message/pixel ξ_i is generated from independent identical distribution (i.i.d.), the probability of the configuration $\{\xi\}$ is written by the product of the probability $P(\xi_i)$, namely, $P(\{\xi\}) = \prod_{i=1}^{N} P(\xi_i)$.

These messages/pixels $\{\xi\}$ are sent through the noisy channel by not only the form $\{\xi_{i1}\cdots\xi_{ip}\} \equiv \{J^0_{i1\cdots ip}\}$ for appropriately chosen set of indexes $\{i1,\cdots,ip\}$ (what we call *parity check* in the context of error-correcting codes) but also sequence of the original messages/pixels itself $\{\xi\}$. Therefore, the outputs of the noisy channel are exchange interactions $\{J_{i1\cdots ip}\}$ and fields $\{\tau_i\}$.

In the field of information theory, the noisy channel is specified by the conditional probability like $P(\{\tau\}|\{\xi\})$ or $P(\{J\}|\{J^0\})$. If each message/pixel ξ_i and parity check $J^0_{i1\cdot ip}$ are affected by the channel noise independently, the probability $P(\{\tau\}|\{\xi\})$ or $P(\{J\}|\{J^0\})$, namely, the probabilities of output sequences $\{\tau\} \equiv (\tau_1,\cdots,\tau_2)$ or $\{J\} \equiv (J_{11\cdots 1p},\cdots,J_{N1\cdots Np})$ for given input sequences $\{\xi\} = (\xi_1,\cdots,\xi_N)$ or $\{J^0\} = (J^0_{11\cdots 1p},\cdots,J^0_{N1\cdots Np})$ are written by

$$P(\{\tau\}|\{\xi\}) = \prod_{i=1}^{N} P(\tau_i|\xi_i), \;\; P(\{J\}|\{J^0\}) = \prod_{i=1}^{N} P(J_{i1\cdots ip}|J^0_{i1\cdots ip}), \qquad (1.1)$$

respectively.

In this chapter, we use the following two kinds of the noisy channel. The first one is referred to as *binary symmetric channel (BSC)*. In this channel, each message/pixel ξ_i and parity check $J_{i1\cdots jp}$ change their sign with probabilities p_τ and p_r, respectively. By introducing the parameters $\beta_\tau \equiv (1/2)\log(1-p_\tau/p_\tau), \beta_r \equiv (1/2)\log(1-p_r/p_r)$, the conditional probabilities (1.1) are given by

$$P(\{\tau\}|\{\xi\}) = \frac{\mathrm{e}^{\beta_\tau \sum_i \tau_i\xi_i}}{[2\cosh\beta_\tau]^N}, \; P(\{J\}|\{J^0\}) = \frac{\mathrm{e}^{\beta_r \sum_{i1\cdots ip} J_{i1\cdots ip}J^0_{i1\cdots ip}}}{[2\cosh\beta_r]^{N_B}} \qquad (1.2)$$

where we defined $N \equiv \sum_i 1, N_B \equiv \sum_{i1\cdots ip} 1$.

Thus, the probability of the output sequences $\{J\}, \{\tau\}$ provided that the corresponding input sequence of the original messages/pixels is $\{\xi\}$ is obtained by $\sum_{\{J_0\}} P(\{J\}|\{J^0\})P(\{J^0\}|\{\xi\})P(\{\tau\}|\{\xi\})$, that is to say,

$$P(\{J\},\{\tau\}|\{\xi\}) = \frac{\exp\left(\beta_r \sum_{i1,\cdots,ip} J_{i1\cdots ip}\,\xi_{i1}\cdots\xi_{ip} + \beta_\tau \sum_i \tau_i \xi_i\right)}{(2\cosh\beta_r)^{N_{\rm B}}(2\cosh\beta_\tau)^N} \tag{1.3}$$

where we used the following condition :

$$P(\{J^0\}|\{\xi\}) = \prod_{i=1}^{N} \delta_{J^0_{i1\cdots ip},\xi_{i1}\cdots\xi_{ip}}. \tag{1.4}$$

The second type of the noisy channel is called as *Gaussian channel (GC)*. The above BSC (1.3) is simply extended to the GC as follows.

$$P(\{J\},\{\tau\}|\{\xi\}) = \frac{\mathrm{e}^{-\frac{1}{2J^2}\sum_{i1,\cdots,ip}(J_{i1\cdots ip}-J_0\xi_{i1}\cdots\xi_{ip})^2 - \frac{1}{2a^2}\sum_i(\tau_i - a_0\xi_i)^2}}{(\sqrt{2\pi}J)^{N_{\rm B}}(\sqrt{2\pi}a)^N}. \tag{1.5}$$

We should notice that these two channels can be treated within the single form :

$$\begin{aligned} P(\{J\},\{\tau\}|\{\xi\}) = &\prod_{i1\cdots ip} F_r(J_{i1\cdots ip}) \prod_i F_\tau(\tau_i) \\ &\times \exp\left(\beta_r \sum_{i1\cdots ip} J_{i1\cdots ip}\,\xi_{i1}\cdots\xi_{ip} + \beta_\tau \sum_i \tau_i\xi_i\right) \end{aligned} \tag{1.6}$$

with

$$F_r(J_{i1\ldots ip}) = \frac{\sum_{j=\pm1}\delta(J_{i1\cdots ip}-j)}{2\cosh\beta_r},\ F_\tau(\tau_i) = \frac{\sum_{j=\pm1}\delta(\tau_i - j)}{2\cosh\beta_\tau} \tag{1.7}$$

for the BSC and

$$F_r(J_{i1\cdots ip}) = \frac{\exp\left[-\frac{1}{2J^2}(J^2_{i1\cdots ip}+J_0^2)\right]}{\sqrt{2\pi J^2}},\ F_\tau(\tau_i) = \frac{\exp\left[-\frac{1}{2a^2}(\tau_i^2+a_0^2)\right]}{\sqrt{2\pi a^2}} \tag{1.8}$$

for the GC. Therefore, it must be noted that there exist relations between the parameters for both channels as

$$\beta_r = \frac{J_0}{J^2},\ \ \beta_\tau = \frac{a_0}{a^2}. \tag{1.9}$$

Main purpose of signal processing we are dealing with in this article is to estimate the original sequence of messages/pixels $\{\xi\}$ from the outputs $\{J\}, \{\tau\}$

of the noisy channel. For this aim, it might be convenient for us to construct the probability of the estimate $\{\sigma\}$ for the original messages/pixels sequence $\{\xi\}$ provided that the outputs of the noisy channel are $\{J\}$ and $\{\tau\}$.

From the Bayes formula, the probability $P(\{\sigma\}|\{J\},\{\tau\})$ is written in terms of the so-called *likelihood* : $P(\{J\},\{\tau\}|\{\sigma\})$ and the *prior* : $P_m(\{\sigma\})$ as follows.

$$P(\{\sigma\}|\{J\},\{\tau\}) = \frac{P(\{J\},\{\tau\}|\{\sigma\})P_m(\{\sigma\})}{\sum_{\{\sigma\}} P(\{J\},\{\tau\}|\{\sigma\})P_m(\{\sigma\})} \tag{1.10}$$

As the likelihood has a meaning of the probabilistic model of the noisy channel, we might choose it naturally as

$$P(\{J\},\{\tau\}|\{\sigma\}) = \frac{\exp\left(\beta_J \sum_{i1\cdots ip} J_{i1\cdots ip}\,\sigma_{i1}\cdots\sigma_{ip} + h\sum_i \tau_i\sigma_i\right)}{(2\cosh\beta_J)^{N_B}(2\cosh h)^N} \tag{1.11}$$

for the BSC and

$$P(\{J\},\{\tau\}|\{\sigma\}) = \frac{\mathrm{e}^{-\frac{\beta_J}{2}\sum_{i1\cdots ip}(J_{i1\cdots ip}-\sigma_{i1}\cdots\sigma_{ip})^2 - h\sum_i(\tau_i-\sigma_i)^2}}{(2\pi/\beta_J)^{N_{\mathrm{B}}/2}(\pi/h)^{N/2}} \tag{1.12}$$

for the GC. Therefore, what we call the *posterior* $P(\{\sigma\}|\{J\},\{\tau\})$ which is defined by (1.10) is rewritten in terms of the above likelihood as follows.

$$P(\{\sigma\}|\{J\},\{\tau\}) = \frac{\mathrm{e}^{-\beta\mathcal{H}_{\mathrm{eff}}}}{\sum_{\{\sigma\}}\mathrm{e}^{-\beta\mathcal{H}_{\mathrm{eff}}}} \tag{1.13}$$

where we defined the inverse temperature $\beta = 1/T$ and set $T = 1$ in the above case. The effective Hamiltonian $\mathcal{H}_{\mathrm{eff}}$ is also defined by

$$\mathcal{H}_{\mathrm{eff}} = -\beta_J \sum_{i1\cdots ip} J_{i1\cdots jp}\,\sigma_{i1}\cdots\sigma_{ip} - h\sum_i \tau_i\sigma_i - \log P_m(\{\sigma\}) \tag{1.14}$$

for the BSC and

$$\mathcal{H}_{\mathrm{eff}} = -\frac{\beta_J}{2}\sum_{i1\cdots ip}(J_{i1\cdots ip}-\sigma_{i1}\cdots\sigma_{ip})^2 - h\sum_i(\tau_i-\sigma_i)^2 - \log P_m(\{\sigma\}) \tag{1.15}$$

for the GC.

1.2.2 MAP estimation and simulated annealing

As we mentioned, the posterior $P(\{\sigma\}|\{J\},\{\tau\})$ is a useful quantity in order to determine the estimate $\{\sigma\}$ of the original messages/pixels sequence. As the estimate of the original message/pixel sequence, we might choose a $\{\sigma\}$ which maximizes the posterior for a given set of the output sequence $\{J\},\{\tau\}$.

Apparently, this estimate $\{\sigma\}$ corresponds to the ground state of the effective Hamiltonian $\mathcal{H}_{\rm eff}$. In the context of Bayesian statistics, this type of estimate $\{\sigma\}$ is referred to as *Maximum A posteriori (MAP)* estimate.

From the view point of important sampling from the posterior as the Gibbs distribution (Gibbs sampling), such a MAP estimate is obtained by controlling the temperature T as $T \to 0$ during the Markov chain Monte Carlo steps. This kind of optimization method is well-known and is widely used as *simulated annealing (SA)* [11, 12]. As the optimal scheduling of the temperature T is $T(t) = c/\log(1+t)$, which was proved by using mathematically rigorous arguments [11].

1.2.3 MPM estimation and a link to statistical mechanics

From the posterior $P(\{\sigma\}|\{J\},\{\tau\})$, we can attempt to make another kind of estimations. For this estimation, we construct the following marginal distribution for each pixel σ_i :

$$P(\sigma_i|\{J\},\{\tau\}) = \sum_{\{\sigma\}\neq\sigma_i} P(\{\sigma\}|\{J\},\{\tau\}). \tag{1.16}$$

Then, we might choose the sign of the difference between $P(1|\{J\},\{\tau\})$ and $P(-1|\{J\},\{\tau\})$ as the estimate of the i-th message/pixel, to put it another way,

$$\mathrm{sgn}\left[\sum_{\sigma_i}\sigma_i P(\sigma_i|\{J\},\{\tau\})\right] = \mathrm{sgn}\left(\frac{\sum_{\{\sigma\}}\sigma_i\,\mathrm{e}^{-\mathcal{H}_{\rm eff}}}{\sum_{\{\sigma\}}\mathrm{e}^{-\mathcal{H}_{\rm eff}}}\right) \equiv \mathrm{sgn}(\langle\sigma_i\rangle_1) \tag{1.17}$$

where we defined the bracket $\langle\cdots\rangle_\beta$ as

$$\langle\cdots\rangle_\beta \equiv \frac{\sum_{\{\sigma\}}(\cdots)\,\mathrm{e}^{-\beta\mathcal{H}_{\rm eff}}}{\sum_{\{\sigma\}}\mathrm{e}^{-\beta\mathcal{H}_{\rm eff}}}. \tag{1.18}$$

Therefore, the above estimate has a link to statistical mechanics through the local magnetization $\langle\sigma_i\rangle_1$ for the spin system that is described by $\mathcal{H}_{\rm eff}$ at temperature $T=1$. This estimate $\mathrm{sgn}(\langle\sigma_i\rangle_1)$ is referred to as *Maximizer of Posterior Marginal (MPM) estimate* or *Finite Temperature (FT) estimate* [6]. It is well-known that this estimate minimizes the following *bit-error rate* :

$$p_b^{(MPM)} = P_b^{(1)}(\beta_J, h : P_m) = \frac{1}{2}\left[1 - R^{(1)}(\beta_J, h : P_m)\right] \tag{1.19}$$

with the overlap between the original message/pixel ξ_i and its MPM estimate $\mathrm{sgn}(\langle\sigma_i\rangle)$:

$$R^{(1)}(\beta_J, h : P_m) = \sum_{\{\xi,J,\tau\}} P(\{J\},\{\tau\},\{\xi\})\,\xi_i\,\mathrm{sgn}(\langle\sigma_i\rangle_1) \tag{1.20}$$

Obviously, the bit-error rate for the MAP estimate is given by

$$p_b^{(MAP)} = \lim_{\beta\to\infty} P^{(\beta)}(\beta_J, h : P_m) = \frac{1}{2}\left[1 - \lim_{\beta\to\infty} R^{(\beta)}(\beta_J, h : P_m)\right] \quad (1.21)$$

with

$$R^{(\beta)}(\beta_J, h : P_m) = \sum_{\{\xi,J,\tau\}} P(\{J\},\{\tau\},\{\xi\})\,\xi_i\,\mathrm{sgn}(\langle\sigma_i\rangle_\beta). \quad (1.22)$$

In the next section, we compare $p_b^{(MPM)}$ with $p_b^{(MAP)}$ by using replica method and show the former is smaller than the later.

1.2.4 The priors and corresponding spin systems

In the previous two subsections, we show the relation between Bayesian inference of the original messages/pixels under some noises and statistical physics [1]. However, we do not yet mention about the choice of the prior distribution $P_m(\{\sigma\})$ in the effective Hamiltonian $\mathcal{H}_{\mathrm{eff}}$. In the framework of the Bayesian statistics, the choice of the prior is arbitrary, however, the quality of the estimation for a given problem strongly depends on the choice.

Image restoration and random field Ising model

In image restoration, we might have an assumption that in the real world two dimensional pictures, the nearest neighboring sites should be inclined to be the same values, in other words, we assume that real picture should be locally smooth (see Fig. 1.1). Taking this smoothness into account, then, it seems reasonable to choose the prior for image restoration as

$$P_m(\{\sigma\}) = \frac{\mathrm{e}^{\beta_m \sum_{<ij>}\sigma_i\sigma_j}}{Z(\beta_m)}, \quad Z(\beta_m) = \sum_{\{\sigma\}} \mathrm{e}^{\beta_m \sum_{<ij>}\sigma_i\sigma_j}. \quad (1.23)$$

In conventional image restoration, we do not send any parity check and only available information is the degraded sequence of the pixels $\{\tau\}$. Thus, we set $\beta_J = 0$ for this problem. Then, we obtain the effective Hamiltonian for image restoration as

$$\mathcal{H}_{\mathrm{eff}} = -\beta_m \sum_{<ij>}\sigma_i\sigma_j - h\sum_i \tau_i\sigma_i \quad (1.24)$$

This Hamiltonian is identical to that of the *random field Ising model* in which random field on each cite corresponds to each degraded pixel τ_i.

Error-correcting codes and spin glasses with p-body interaction

In error-correcting codes, we usually use so-called *uniform distribution* because we do not have any idea about the properties of the original message sequence $\{\xi\}$ as we assumed smoothness for images. Thus, we set the prior as $P_m(\{\sigma\}) = 2^{-N}$ and substituting $-\log P_m(\{\sigma\}) = N \log 2 = \text{const.}$ into $\mathcal{H}_{\text{eff}}$ (usually, we neglect the constant term).

In this case, we do not use any a priori information to estimate the original message, however, in error-correcting codes, we compensate this lack of information with extra redundant information as a form of $\xi_{i1} \cdots \xi_{ip}$, besides the original message sequence $\{\xi\}$. In information theory, it is well-known that we can decode the original message $\{\xi\}$ without any error when the transmission rate R, which is defined by $R = N/N_B$ (N original message length, N_B : redundant message length), is smaller than the channel capacity C (see for example [26]). The channel capacity is given by $C = 1+p\log_2 p+(1-p)\log_2(1-p)$ for the BSC and $p = (1/2)\log_2(1+J_0^2/J^2)$ for the GC. As we will mention in the next section, when we send ${}_NC_r$ combinations of p bits among the original image $\{\xi\}$, as products $\xi_{i1} \cdots \xi_{ip}$, error-less decoding might be achieved in the limit of $p \to \infty$. We call this type of code as *Sourlas codes* [5]. For this Sourlas codes, we obtain the following effective Hamiltonian.

$$\mathcal{H}_{\text{eff}} = -\beta_J \sum_{i1\cdots ip} J_{i1\cdots ip}\, \sigma_{i1} \cdots \sigma_{ip} - h \sum_i \tau_i \sigma_i \tag{1.25}$$

It is clear that this Hamiltonian is identical to that of the *Ising spin glass model with p-body interaction* under some random fields on cites.

1.3 Quantum version of the model

In the previous sections, we explained the relation between the Bayesian statistics and statistical mechanics. We found that there exists the effective Hamiltonian for each problem of image restoration and error-correcting codes. In order to extend the model systems to their quantum version, we add the transverse field term : $-\Gamma \sum_i \sigma_i^x$ into the effective Hamiltonian [22]. In this expression, $\{\sigma^x\}$ means the x-component of the Pauli matrix and Γ controls the strength of *quantum fluctuation*. Each term $\Gamma\sigma_i^x$ appearing in the sum might be understood as tunneling probability between the states $\sigma_i^z = +1$ and $\sigma_i^z = -1$ intuitively. As the result, the quantum version of image restoration is reduced to that of statistical mechanics for the following effective Hamiltonian

$$\mathcal{H}_{\text{eff}}^{Quantum} = -\beta_m \sum_{<ij>} \sigma_i^z \sigma_j^z - h \sum_i \tau_i \sigma_i^z - \Gamma \sum_i \sigma_i^x. \tag{1.26}$$

We also obtain the quantum version of the effective Hamiltonian for error-correcting codes as

$$\mathcal{H}_{\rm eff}^{Quantum} = -\beta_J \sum_{i1\cdots ip} J_{i1\cdots ip}\, \sigma_{i1}^z \cdots \sigma_{ip}^z - h \sum_i \tau_i \sigma_i^z - \Gamma \sum_i \sigma_i^x. \quad (1.27)$$

We should keep in mind that in the context of the MAP estimation, it might be useful for us to controlling the strength of the quantum fluctuation, namely, the amplitude of the transverse field Γ as $\Gamma \to 0$ during the quantum Markov chain Monte Carlo steps. If this annealing process of Γ is slow enough, at the end $\Gamma = 0$, we might obtain the ground states of the classical spin systems described by the following Hamiltonian

$$\mathcal{H}_{\rm eff}^{classical} = -\beta_m \sum_{<ij>} \sigma_i^z \sigma_j^z - h \sum_i \tau_i \sigma_i^z \quad (1.28)$$

for image restoration and

$$\mathcal{H}_{\rm eff}^{classical} = -\beta_J \sum_{i1\cdots ip} J_{i1\cdots ip}\, \sigma_{i1}^z \cdots \sigma_{ip}^z - h \sum_i \tau_i \sigma_i^z. \quad (1.29)$$

for error-correcting codes. This is an essential idea of the *quantum annealing*. Unfortunately, up to now, there are no mathematically rigorous arguments for the optimal scheduling of $\Gamma(t)$ corresponds to Geman and Geman's proofs [11] for the simulated annealing [12]. We will revisit this problem in last section of this article. In this article, we investigate its averaged case performance by analysis of the infinite range model and by caring out quantum Markov chain Monte Carlo simulations.

1.4 Analysis of the infinite range model

In the previous section, we completely defined our two problems of information processing, that is to say, image restoration and error-correcting codes as the problems of statistical mechanics of random spin systems in a transverse field. We found that there exist two possible candidates to determine the original sequence of the messages/pixels. The first one is the MAP estimation and the estimate is regarded as ground states of the effective Hamiltonian that is defined as a minus of logarithm of the posterior distribution. As we mentioned, to carry out the optimization of the Hamiltonian, both the simulated annealing and the quantum annealing are applicable. In order to construct the quantum annealing, we should add the transverse field to the effective Hamiltonian and control the amplitude of the field Γ during the quantum Markov chain Monte Carlo steps. Therefore, the possible extension of the classical spin systems to the corresponding quantum spin systems in terms of the transverse field is essential idea of our work.

Besides the MAP estimate as a solution of the optimization problems, the MPM estimate, which is given by the sign of the local magnetization of the spin system, is also available. This estimate is well-known as the estimate

that minimizes the bit-error rate. Performances of both the MAP and the MPM estimations are evaluated through this bit-error rate.

In order to evaluate the performance, we first attempt to calculate the bit-error rate analytically by using the mean-field infinite range model. As the most famous example of solvable model, Sherrington-Kirkpatrick model [9] in spin glasses, we also introduce the solvable models for both image restoration and error-correcting codes. In this section, according to the previous work by the present author [21], we first investigate the performance of image restoration.

It is important to bear in mind that in our Hamiltonian, there exists two types of terms, namely, $\mathcal{A}_0 = -\mathcal{H}_{\rm eff}^{classical}$ and $\mathcal{A}_1 = -\Gamma \sum_i \sigma_i^x$, and they do not commute with each other. Therefore, it is impossible to calculate the partition function directly. Then, we use the *Suzuki-Trotter (ST) decomposition* [24, 25]

$$Z_{\rm eff} = \lim_{M\to\infty} {\rm tr} \left({\rm e}^{\frac{\mathcal{A}_0}{M}} {\rm e}^{\frac{\mathcal{A}_1}{M}} \right)^M \tag{1.30}$$

to cast the problem into an equivalent classical spin system. In following, we calculate the macroscopic behavior of the model system with the assistance of the ST formula [24, 25] and replica method [9] for the data $\{\xi, J, \tau\}$ average $[\cdots]_{data}$:

$$[\log Z_{\rm eff}]_{data} = \lim_{n\to 0} \frac{[Z_{\rm eff}^n]_{data} - 1}{n} \tag{1.31}$$

of the infinite range model.

1.4.1 Image restoration

In order to analyze the performance of the MAP and the MPM estimation in image restoration, we suppose that the original image is generated by the next probability distribution,

$$P(\{\xi\}) = \frac{{\rm e}^{\frac{\beta_s}{N}\sum_{ij}\xi_i\xi_j}}{Z(\beta_s)}, \quad Z(\beta_s) = \sum_{\{\xi\}} {\rm e}^{\frac{\beta_s}{N}\sum_{ij}\xi_i\xi_j}, \tag{1.32}$$

namely, the Gibbs distribution of the ferromagnetic Ising model at the temperature $T_s = \beta_s^{-1}$. For this original image and under the Gaussian channel, the macroscopic properties of the system like the bit-error rate are derived from the data-averaged free energy $[\log Z_{\rm eff}]_{data}$. Using the ST formula and the replica method, we write down the replicated partition function as follows.

$$\begin{aligned}[Z_{\rm eff}^n]_{data} = \sum_{\{\xi\}} \int_{-\infty}^{\infty} \prod_{ij} \frac{dJ_{ij}}{\sqrt{2\pi J^2/N}} {\rm e}^{-\frac{N}{2J^2}\sum_{ij}\left(J_{ij} - \frac{J_0}{N}\xi_i\xi_j\right)^2} \\ \times \int_{-\infty}^{\infty} \prod_i \frac{d\tau_i}{\sqrt{2\pi}a} {\rm e}^{-\frac{1}{2a^2}\sum_i(\tau_i - a_0\xi_i)^2} \times \frac{{\rm e}^{(\beta_s/N)\sum_{ij}\xi_i\xi_j}}{Z(\beta_s)}\end{aligned}$$

$$\times \mathrm{tr}_{\{\sigma\}} \prod_{\alpha=1}^{n} \prod_{K=1}^{M} \exp\left[\frac{\beta_J}{M} \sum_{ij} J_{ij} \sigma_{iK}^{\alpha} \sigma_{jK}^{\alpha} + \frac{\beta_m}{MN} \sum_{ij} \sigma_{iK}^{\alpha} \sigma_{jK}^{\alpha} \right.$$
$$\left. + \frac{h}{M} \sum_{i} \tau_i \sigma_{iK}^{\alpha} + B \sum_{i} \sigma_{iK}^{\alpha} \sigma_{i,K+1}^{\alpha} \right] \tag{1.33}$$

where $[\cdots]_{data}$ means average over the quenched randomness, namely, over the joint probability $P(\{J\},\{\tau\},\{\xi\})$. We should keep in mind that these quantities $\{\xi\}$ and $\{J\},\{\tau\}$ mean the data we send to the receiver and the outputs of the channel the receiver obtain, respectively. Therefore, by calculating these averages $[\cdots]_{data}$, we can evaluate the *data-averaged case performance* of the image restoration [21]. We also defined the partition function $Z(\beta_s)$ for the original images and B as $Z(\beta_s) \equiv \sum_{\{\xi\}} \mathrm{e}^{(\beta_s/N)\sum_{ij}\xi_i\xi_j}, B \equiv (1/2)\log\coth(\Gamma/M)$. The standard replica calculation leads to the following expressions of the free energy density :

$$[\log Z_{\mathrm{eff}}]_{data} = \frac{[Z_{\mathrm{eff}}^{n}]_{data} - 1}{nN} = -\frac{f_0^{RS}}{n} - f^{RS} \tag{1.34}$$

$$f_0^{RS} = \frac{1}{2}\beta_s m_0^2 - \log 2\cosh(\beta_s m_0) \tag{1.35}$$

$$f^{RS} = -\frac{(\beta_J J)^2}{2} Q^2 + \frac{(\beta_J J)^2}{2} S^2 + \frac{\beta_m}{2} m^2 + \frac{\beta_J J_0}{2} t^2$$
$$- \sum_{\xi} \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \log \int_{-\infty}^{\infty} Dw\, 2\cosh\sqrt{\Phi^2 + \Gamma^2} \tag{1.36}$$

and the saddle point equations with respect to the order parameters.

$$[\langle \sigma_{iK}^{\alpha} \rangle]_{data} = m = \sum_{\xi} \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \int_{-\infty}^{\infty} D\omega \left(\frac{\Phi \sinh \Xi}{\Xi \Omega} \right) \tag{1.37}$$

$$[\xi_i \langle \sigma_{iK}^{\alpha} \rangle]_{data} = t = \sum_{\xi} \xi \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \int_{-\infty}^{\infty} D\omega \left(\frac{\Phi \sinh \Xi}{\Xi \Omega} \right) \tag{1.38}$$

$$[\langle (\sigma_{iK}^{\alpha})^2 \rangle]_{data} = Q = \sum_{\xi} \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \left[\int_{-\infty}^{\infty} D\omega \left(\frac{\Phi \sinh \Xi}{\Xi \Omega} \right) \right]^2 \tag{1.39}$$

$$[\langle \sigma_{iK}^{\alpha} \sigma_{iL}^{\alpha} \rangle]_{data} = S$$
$$= \sum_{\xi} \mathcal{M}(\xi) \int_{-\infty}^{\infty} \frac{Du}{\Omega} \int_{-\infty}^{\infty} \left[\left(\frac{\Phi}{\Xi} \right)^2 \cosh \Xi + \Gamma^2 \left(\frac{\sinh \Xi}{\Xi^3} \right) \right] \tag{1.40}$$

with $[\xi_i]_{data} = m_0 = \tanh(\beta_s m_0)$ and $\mathcal{M}(\xi) = \mathrm{e}^{\beta_s m_0 \xi}/2\cosh(\beta_s m_0)$. where we used the replica symmetric and the static approximation, that is,

$$t_K = t,\ \ S_\alpha(KL) = \begin{cases} S\ (K \neq L) \\ 1\ \ (K = L) \end{cases},\ \ Q_{\alpha\beta} = Q \tag{1.41}$$

and $\langle \cdots \rangle$ denotes the average over the posterior distribution and Φ, y and Ω are defined as

$$\Phi \equiv u\sqrt{(ah)^2 + (J\beta_J)^2 Q} + J\beta_J \omega \sqrt{S - Q} + (a_0 h + J_0 \beta_J t)\xi + \beta_m m \tag{1.42}$$

$$\Xi \equiv \sqrt{\Phi^2 + \Gamma^2},\ \ \Omega \equiv \int_{-\infty}^{\infty} D\omega\ \cosh\Xi. \tag{1.43}$$

Then, the overlap R which is a measure of retrieval quality is calculated explicitly as

$$[\xi_i \operatorname{sgn}(\langle \sigma_{iK}^{\alpha} \rangle)]_{data} = R = \sum_{\xi} \xi \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \int_{-\infty}^{\infty} Dw \operatorname{sgn}(\Phi), \tag{1.44}$$

then, of course, the bit-error rate is given by $p_b = (1 - R)/2$.

1.4.2 Image restoration at finite temperature

We first investigate the image restoration without parity check term $\beta_J = 0$. For this case, the saddle point equations lead to the following much simpler coupled equations :

$$m_0 = \tanh(\beta_s m_0),\ \ m = \sum_{\xi} \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du\, \frac{\Phi_0\ \tanh\sqrt{\Phi_0^2 + \Gamma^2}}{\sqrt{\Phi_0^2 + \Gamma^2}} \tag{1.45}$$

with $\Phi_0 \equiv m\beta_m + a_0 h\xi + ahu$. Then, the overlap R is also reduced to

$$R = \sum_{\xi} \xi \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \operatorname{sgn}(\Phi_0) = 1 - 2p_b \tag{1.46}$$

where R depends on Γ through m. In Fig. 1.2 (left), for the case of no parity check $\beta_J = 0$, we plot the bit-error rate p_b as a function of $T_m = \beta_m^{-1}$. We choose the temperature of the original image $T_s^{-1} = \beta_s = 0.9$ and noise rate $\beta_\tau = a_0/a^2 = 1$. We keep the ratio h/β_m to its optimal value $\beta_\tau/\beta_s = 0.9$ and investigate T_m-dependence of p_b. Then, the parameter T_m has a meaning of temperature for simulated annealing. Obviously, $p_b^{(MAP)} = \lim_{T_m \to 0} p_b$ and p_b at $T_m = T_s$ is the lowest value of $p_b^{(MPM)}$ for $\Gamma = 0$.

Let us stress again that in practice, the infinite range model is not useful for realistic two dimensional image restoration because all pixels are neighbor each other. In order to restore these two dimensional images, we should use the prior $P(\{\xi\})$ for two dimension. In fact, let us think about the overlap r between an original pixel ξ_i and corresponding degraded pixel τ_i, namely,

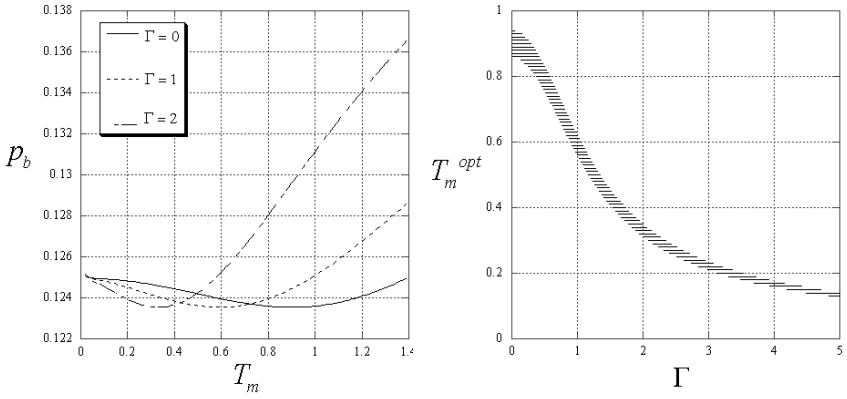

Fig. 1.2. The bit-error rate $p_b = (1 - R)/2$ without exchange term ($\beta_J = 0$) as a function of temperature $T_m = \beta_m^{-1}$ (left). Keeping the ratio to $h/\beta_m = \beta_\tau/\beta_s = a_0/a^2\beta_s = 0.9$ (we set $a_0 = a = 1$), we change the value of T_m. For the case of $\Gamma = 0$, p_b takes its minimum at $T_m = T_s = 0.9$. For finite Γ, the optimal temperature T_m is not T_s, however, the minimum of p_b does not change. The right panel shows the optimal temperature T_m^{opt} as a function of Γ.

$$r = [\xi_i \tau_i]_{data} = \frac{\sum_{\tau,\xi} \mathrm{e}^{\beta_\tau \xi\tau + \beta_s m_0 \xi}(\xi\tau)}{4\cosh(\beta_\tau)\cosh(\beta_s m_0)} = \tanh(\beta_\tau). \tag{1.47}$$

From this relation, the error probability p_τ is given as $p_\tau = (1 - r)/2 = 1/(1 + \mathrm{e}^{2\beta_\tau}) = 0.119 < p_b^{(MPM)}$ for $\beta_\tau = 1$, and unfortunately, the restored image becomes much worse than the degraded (see Fig. 1.2 (left)). This is because any spacial structure is ignored in this artificial model. This result might be understood as a situation in which we try to restore the finite dimensional image with some structures by using the infinite range prior without any structure (namely, the correlation length between pixels is also infinite). However, the infinite range model is useful to predict the qualitative behavior of macroscopic quantities like bit-error rate and we can grasp the details of its hyperparameters (namely, T_m, h or Γ) dependence and can also compare the MAP with the MPM estimations. This is a reason why we introduce this model to the analysis of image restoration problems. Of course, if we use two dimensional structural priors, the both the MAP estimations via simulated and quantum annealing and the MPM estimation by using thermal and quantum fluctuations work well for realistic two dimensional image restoration. In the next section, we will revisit this problem and find it. It is also important for us to bear in mind that the quality of the restoration depends on the macroscopic properties of the original image.

In our choice of the original image, its macroscopic qualities are determined by the temperature T_s and magnetization m_0 as a solution of $m_0 = \tanh(\beta_s m_0)$. Although we chose the temperature $T_s = 0.9$ in Fig. 1.2 (left), it is important to check the retrieval quality for different temperatures T_s. In Fig. 1.3 (right), we plot the bit-error rate for the case of $T_s = 0.7$. From this panel, we find $p_b < p_\tau$ and the MPM estimation improves the quality of the restoration.

For $\Gamma > 0$, the optimal temperature which gives the minimum of p_b is not T_s. In the right panel of Fig. 1.2, we plot the $T_m^{\rm opt}$ as a function of Γ. In Fig. 1.3

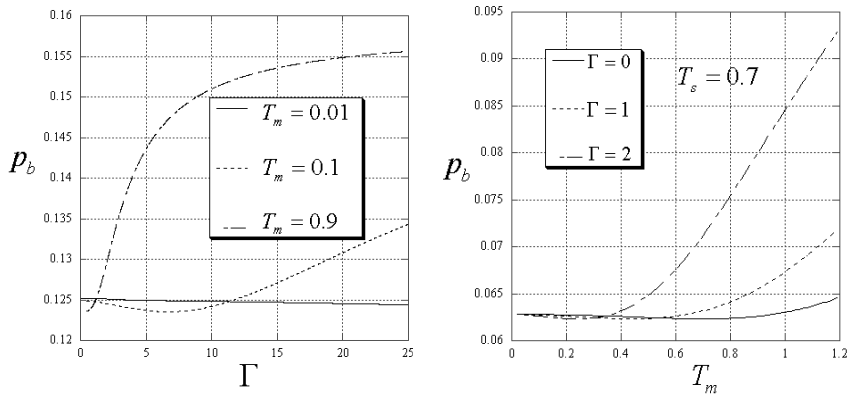

Fig. 1.3. The bit-error rate p_b is drawn for $T_m = 0.01, 0.1$ and $T_m = 0.9$ as a function of Γ (left). The right panel is the same type of the plot as the right panel in Fig. 1.2 for the case of $T_s = 0.7$.

(left), we plot the bit-error rate as a function of Γ for $T_m = T_s = 0.9$ setting the ratio to its optimal value $h/\beta_m = \beta_\tau/\beta_s = 0.9$. From this figure, we find that the MPM optimal estimate no longer exists by adding the transverse field $\Gamma > 0$ and the bit-error rate p_b increases as the amplitude of the transverse field Γ becomes much stronger.

On the other hand, when we set the temperature $T_m = 0.01$, the Γ-dependence of the bit-error rate is almost flat (see Fig. 1.3 (right)). We should notice that p_b at $\Gamma = 0$ for $T_m = 0$ corresponds to the performance of the MAP estimation by quantum annealing. We discuss the performance of the quantum annealing in the last part of this subsection.

We next consider the performance for the MAP and the MPM estimations with parity check term ($\beta_J \neq 0$). We plot the result in FIG. 1.4. As we mentioned before, two body parity check term works very well to decrease the bit-error rate p_b. However, in this case, there does not exist the optimal

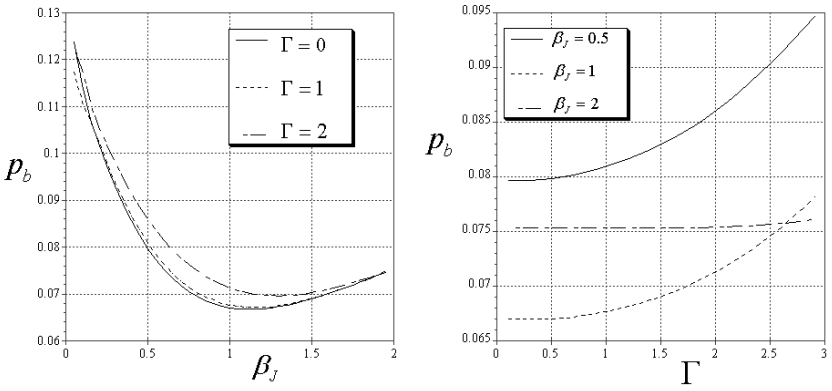

Fig. 1.4. The bit-error rate p_b as a function of β_J for $\Gamma = 0, 1, 2$ keeping the ration constant $h/\beta_m = \beta_\tau/\beta_s$ (left). In right panel, p_b as a function of Γ is plotted for the case of $\beta_J = 0.5, 1, 2.0$.

β_J which minimizes the bit-error rate for any finite values of Γ. As we see the left panel in FIG. 1.4, for small value of β_J, the restoration by a finite Γ is superior to that of absence of the transverse field ($\Gamma = 0$).

Hyperparameter estimation

In this subsection, we evaluated the performance of the MAP and the MPM estimations in image restoration through the bit-error rate. In these results, we found that the macroscopic parameters, β_m, h and Γ-dependence of the bit-error rate have important information to retrieve the original image. However, from the definition, (1.44)(1.46), as the bit-error rate contains the original image $\{\xi\}$, it is impossible for us to use p_b as a cost function to determine the best choice of these parameters. In statistics, we usually use the *marginal likelihood* [27] which is defined by the logarithm of the normalization constant of $\mathrm{tr}_{\{\sigma\}} P(\{\sigma\}|\{\tau\})P_m(\{\sigma\})$, that is,

$$K(\beta_m, h, \Gamma : \{\tau\}) \equiv \log Z_{Pos.} - \log Z_{Pri.} - \log Z_L \tag{1.48}$$

where $Z_{Pos.}, Z_{Pri}$ and Z_L are normalization constants for the posterior, the prior and the likelihood, and which are given by

$$Z_{Pos.} = \mathrm{tr}_{\{\sigma\}} \mathrm{e}^{\beta_m \sum_{ij} \sigma_i^z \sigma_j^z + h \sum_i \tau_i \sigma_i^z + \Gamma \sum_i \sigma_i^x} \tag{1.49}$$

$$Z_{Pri.} = \mathrm{tr}_{\{\sigma\}} \mathrm{e}^{\beta_m \sum_{ij} \sigma_i^z \sigma_j^z + \Gamma \sum_i \sigma_i^x}, \quad Z_L = \mathrm{tr}_{\{\tau\}} \mathrm{e}^{h \sum_i \tau_i \sigma_i^z}, \tag{1.50}$$

respectively. For simplicity, let us concentrate ourself to the case of no parity check $\beta_J = 0$.

It must be noted that the marginal likelihood (1.48) is constructed by using the observables $\{\tau\}$ and does not contain the original image $\{\xi\}$ at all. Therefore, in practice, the marginal likelihood has a lot of information to determine the macroscopic parameters, what we call *hyperparameters*, before we calculate the MAP and the MPM estimates.

In the infinite range model, it is possible for us to derive the data-averaged marginal likelihood per pixel $K(\beta_J, h, \Gamma) = [K(\beta_J, h, \Gamma : \{\tau\})]_{data}/N$ explicitly. Here we first investigate the hyperparameter dependence of the marginal likelihood. $\log Z_{Pri}$ and $[\log Z_L]_{data} = [\log \int_{-\infty}^{\infty} \prod_i d\tau_i F_\tau(\tau_i) \mathrm{e}^{h\tau_i \sigma_i^z}]_{data}$ per pixel can be calculated as

$$\frac{\log Z_{Pri}}{N} = -\frac{\beta_m m_1^2}{2} + \log 2\cosh\sqrt{(\beta_m m_1)^2 + \Gamma^2} \tag{1.51}$$

$$\frac{[\log Z_L]_{data}}{N} = -\frac{h^2}{2}\left[\left(\frac{a_0}{ah}\right)^2 - a^2\right] \tag{1.52}$$

and the data average of the first term of the right hand side of (1.48) is identical to the free energy density for $\beta_J = 0$. Thus, we obtain the data-averaged marginal likelihood as follows.

$$K(\beta_m, h, \Gamma) = -\frac{\beta_m m^2}{2} + \sum_{\xi} \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du \log 2\cosh\sqrt{\Phi_0^2 + \Gamma^2}$$
$$+ \frac{\beta_m m_1^2}{2} - \log 2\cosh\sqrt{(\beta_m m_1)^2 + \Gamma^2} + \frac{h^2}{2}\left[\left(\frac{a_0}{ah}\right)^2 - a^2\right] \tag{1.53}$$

where m_1, m mean the magnetizations of the prior and the posterior, are given by

$$m_1 = \frac{\beta_m m_1 \tanh\sqrt{(\beta_m m_1)^2 + \Gamma^2}}{\sqrt{(\beta_m m_1)^2 + \Gamma^2}} \tag{1.54}$$

and (1.45), respectively. In Fig. 1.5 (left), we plot $K(\beta_m, h, \Gamma)$. In this figure, we set $T_s = 0.9, \beta_\tau = 1$. We found that the data-averaged marginal likelihood takes its maximum at $T_m = T_s, h = \beta_\tau$ and $\Gamma = 0$. This result might be naturally understood because the performance of both the MAP and MPM estimation should be the best for setting the probabilistic models of the noise channel and the distribution of the original image to the corresponding true probabilities. Therefore, it might seems that the transverse field Γ has no meaning for restoration. However, when we attempt to maximize the marginal likelihood via gradient descent, we need to solve the following coupled equations.

$$c_{\beta_m}\frac{d\beta_m}{dt} = \frac{\partial K}{\partial \beta_m} = \langle \sum_{ij} \sigma_i^z \sigma_j^z \rangle_{Pos.} - \langle \sum_{ij} \sigma_i^z \sigma_j^z \rangle_{Pri.} \tag{1.55}$$

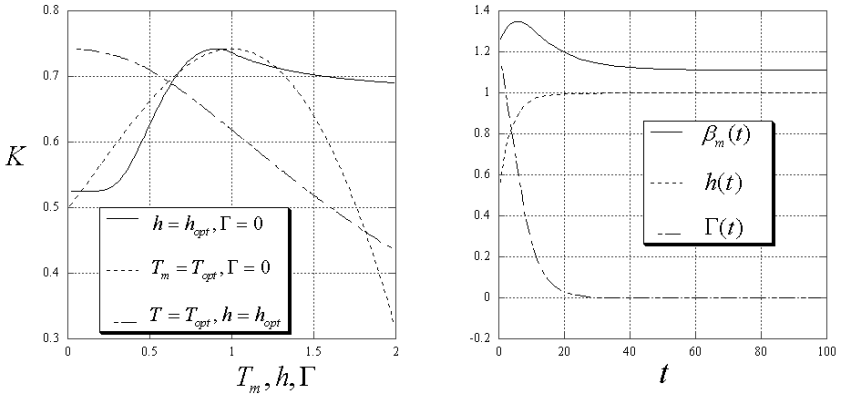

Fig. 1.5. The data-averaged marginal as a function of hyperparameters, β_m, h and Γ (left). The right panel shows the time development of the hyperparameter β_m, h and Γ via gradient descent of the marginal likelihood. We set the time constants $c_{\beta_m} = c_h = c_\Gamma = 1$ and the values of true hyperparameters as $T_s = \beta_s^{-1} = 0.9, \beta_\tau = 1$.

$$c_h \frac{dh}{dt} = \frac{\partial K}{\partial h} = \langle \sum_i \tau_i \sigma_i^z \rangle_{Pos.} - \langle \sum_i \tau_i \sigma_i^z \rangle_{Pri.} - \langle \sum_i \tau_i \sigma_i^z \rangle_L \quad (1.56)$$

$$c_\Gamma \frac{d\Gamma}{dt} = \frac{\partial K}{\partial \Gamma} = \langle \sum_i \sigma_i^x \rangle_{Pos.} - \langle \sum_i \sigma_i^x \rangle_{Pri.} \quad (1.57)$$

with the definitions of the brackets

$$\langle \cdots \rangle_{Pos.} = \frac{\mathrm{tr}_{\{\sigma\}} (\cdots) \, \mathrm{e}^{\beta_m \sum_{ij} \sigma_i^z \sigma_j^z + h \sum_i \tau_i \sigma_i^z + \Gamma \sum_i \sigma_i^x}}{\mathrm{tr}_{\{\sigma\}} \mathrm{e}^{\beta_m \sum_{ij} \sigma_i^z \sigma_j^z + h \sum_i \tau_i \sigma_i^z + \Gamma \sum_i \sigma_i^x}} \quad (1.58)$$

$$\langle \cdots \rangle_{Pri.} = \frac{\mathrm{tr}_{\{\sigma\}} (\cdots) \, \mathrm{e}^{\beta_m \sum_{ij} \sigma_i^z \sigma_j^z + \Gamma \sum_i \sigma_i^x}}{\mathrm{tr}_{\{\sigma\}} \mathrm{e}^{\beta_m \sum_{ij} \sigma_i^z \sigma_j^z + \Gamma \sum_i \sigma_i^x}}, \quad \langle \cdots \rangle_L = \frac{\mathrm{tr}_{\{\tau\}} (\cdots) \, \mathrm{e}^{h \sum_i \tau_i \sigma_i^x}}{\mathrm{tr}_{\{\tau\}} \mathrm{e}^{h \sum_i \tau_i \sigma_i^x}} \quad (1.59)$$

and time constants c_{β_m}, c_h and c_Γ. Thus, when we solve the above equations, we need to evaluate these expectations for every time steps by using the quantum Markov chain Monte Carlo method. It is obvious that it takes quite long time to obtain the solutions. From reasons mentioned above, it is convenient for us to suppress the error of hyperparameter estimation by introducing the transverse field. From figures, Fig. 1.2, Fig. 1.3, we actually find these desirable properties.

Incidentally, for the infinite range model, we derive these coupled equations explicitly. The results are given by

$$c_{\beta_m}\frac{d\beta_m}{dt} = \frac{m_1^2 - m^2}{2} - \frac{\beta_m m_1^2 \tanh\sqrt{(\beta_m m_1)^2 + \Gamma^2}}{\sqrt{(\beta_m m_1)^2 + \Gamma^2}} + m\sum_{\xi}\mathcal{M}(\xi)\int_{-\infty}^{\infty} Du\,\frac{\Phi_0 \tanh\sqrt{\Phi_0^2 + \Gamma^2}}{\sqrt{\Phi_0^2 + \Gamma^2}} \tag{1.60}$$

$$c_h\frac{dh}{dt} = -a^2 h + \sum_{\xi}\mathcal{M}(\xi)\int_{-\infty}^{\infty} Du\,\frac{\Phi_0(a_0\xi + au)\tanh\sqrt{\Phi_0^2 + \Gamma^2}}{\sqrt{\Phi_0^2 + \Gamma^2}} \tag{1.61}$$

$$c_\Gamma\frac{d\Gamma}{dt} = -\frac{\Gamma\tanh\sqrt{(\beta_m m_1)^2 + \Gamma^2}}{\sqrt{(\beta_m m_1)^2 + \Gamma^2}} + \Gamma\sum_{\xi}\mathcal{M}(\xi)\int_{-\infty}^{\infty} Du\,\frac{\tanh\sqrt{\Phi_0^2 + \Gamma^2}}{\sqrt{\Phi_0^2 + \Gamma^2}} \tag{1.62}$$

where m_1 and m satisfy (1.45) and (1.54). We plot the results by solving the differential equations with respect to the hyperparameters, namely, (1.60)(1.61)(1.62) numerically in Fig. 1.5. We find that each hyperparameter converges to its optimal value.

1.4.3 Image restoration driven by pure quantum fluctuation

In the above discussion, we investigated mainly the MPM estimation at finite temperature $T_m > 0$ according to the reference [21]. However, it is worth while for us to check the following limit : $\beta_m \to \infty$ keeping the *effective amplitude of transverse field* $\Gamma_{\rm eff} = \Gamma/\beta_m$ finite. In this limit, we investigate pure effect of the quantum fluctuation without any thermal one. To evaluate the performances of the MAP and the MPM estimations for this zero temperature case, we set $\Phi_0 = \beta_m(m + h_* a_0\xi + h_* a u) = \beta_m\phi_0$, where h_* is its optimal value $h_* = \beta_s/\beta_\tau$, and consider the asymptotic form of the saddle point equations with respect to m and m_1 in the limit of $\beta_m \to \infty$. We easily find

$$m_1 = \sqrt{1 - \Gamma_{\rm eff}^2},\ \ m = \sum_{\xi}\mathcal{M}(\xi)\int_{-\infty}^{\infty}\frac{\phi_0 Du}{\sqrt{\phi_0^2 + \Gamma_{\rm eff}^2}} \tag{1.63}$$

and the time evolution of $\Gamma_{\rm eff}$ as follows.

$$c_{\Gamma_{\rm eff}}\frac{d\Gamma_{\rm eff}}{dt} = -\frac{\Gamma_{\rm eff}}{\sqrt{m_1^2 + \Gamma_{\rm eff}^2}} + \sum_{\xi}\mathcal{M}(\xi)\int_{-\infty}^{\infty}\frac{\Gamma_{\rm eff} Du}{\sqrt{\phi_0^2 + \Gamma_{\rm eff}^2}} \tag{1.64}$$

where $c_{\Gamma_{\rm eff}} = \beta_m c_\Gamma$. The bit-error rate is given by $p_b = (1 - m_0)/2 + \sum_\xi \mathcal{M}(\xi)\xi H(u_*)$, where $u_* = (a_0 h_*\xi + m)/a h_*$. We fist plot the $\Gamma_{\rm eff}$-dependence of the bit-error rate at $T_m = 0$ in Fig. 1.6. In this figure, the value at $\Gamma_{\rm eff} = 0$ corresponds to the quantum MAP estimation which might be realized by the quantum annealing. From this figure, we find that the

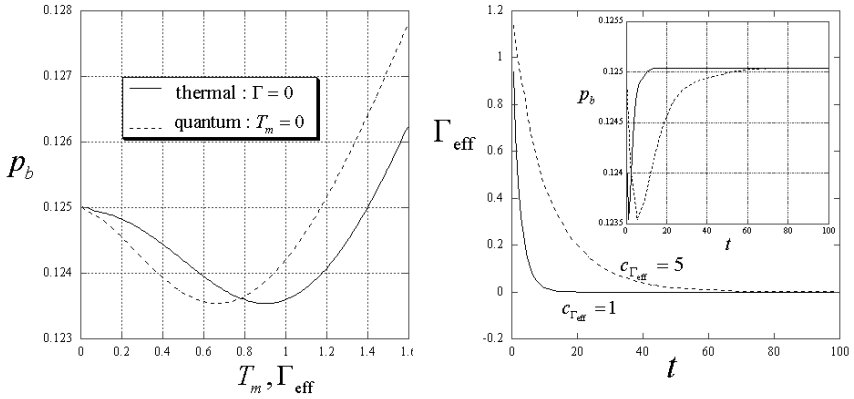

Fig. 1.6. The bit-error rate for the quantum ($T_m = 0$) and the thermal (classical) ($\Gamma = 0$) estimation (left). The right panel shows the time development of the effective amplitude of the transverse field $\Gamma_{\rm eff} = \Gamma/\beta_m$. The inset means the time dependence of the bit-error rate

performance of the quantum MPM estimation is superior to the MAP estimation and there exists some finite value of the amplitude Γ at which the bit-error rate takes its minimum. In the same figure, we also plot the T_m-dependence of the bit-error rate for $\Gamma = 0$. We find that ,for both the quantum and the thermal cases, the best possible values of both the MAP and the MPM estimation is exactly the same. In Fig.1.6 (right), we plot the time development of the effective amplitude of transverse field and the resultant bit-error rate. From this figure, we notice that at the beginning of the gradient descent the bit-error rate decreases but as Γ decreases to zero, the error converges to the best possible value for the quantum MAP estimation. The speed of the convergence is exponentially fast. Actually, in the asymptotic limit $t \to \infty, \Gamma_{\rm eff} \to 0$, the equation (1.64) is solved as $\Gamma_{\rm eff} = \Gamma_{\rm eff}(0)\,{\rm e}^{-\theta_{\Gamma_{\rm eff}} t}$, where $\theta_{\Gamma_{\rm eff}} \equiv (1/c_{\Gamma_{\rm eff}})(1 - \sum_\xi \mathcal{M}(\xi) \int_{-\infty}^{\infty} Du/|\phi_0|)$. However, this fact does not mean that it is possible for us to decrease the effective amplitude of the transverse field to zero by using exponentially fast scheduling to realize the best possible performance of the quantum MAP estimation. This is because the time unit t appearing in (1.64) does not corresponds to the quantum Monte Carlo step and the dynamics (1.64) requires the (equilibrium) magnetization $m(\Gamma_{\rm eff})$ at each time step in the differential equation. As the result, we need the information about m near $\Gamma_{\rm eff} \to 0$, namely, the asymptotic form : $m(t \to \infty, \Gamma_{\rm eff} \to 0)$ to discuss the annealing schedule to obtain the MAP estimation. Although we assume that each time step in (1.64), the system obeys the equilibrium condition : $m = \sum_\xi \mathcal{M}(\xi) \int_{-\infty}^{\infty} \phi_0 Du/\sqrt{\phi_0^2 + \Gamma_{\rm eff}^2}$, we

need the dynamics of m to discuss the optimal annealing scheduling about $\Gamma_{\rm eff}$. This point will be discussed in last section by means of the quantum Markov chain Monte Carlo method.

The Nishimori-Wong condition on the effective transverse field

From Fig. 1.6 (left), we found that the lowest value of the bit-error rate is same both for the thermal and the quantum MPM estimations. In the thermal MPM estimation, Nishimori and Wong [8] found that the condition on which the best performance is obtained, namely, what we call *Nishimori-Wong condition*. They showed that the condition : $(m/m_0) = (h/\beta_\tau)(\beta_s/\beta_m)$ should hold in order to obtain the lowest value of the bit-error rate. When we set the hyperparameter h to its true value $h = \beta_\tau$, the condition is reduced to the simple form : $T_m^{\rm opt} = T_s$. Therefore, it is important for us to derive the same kind of condition which gives the best performance of the quantum MPM estimation. Here we derive the condition and show the lowest values of the p_b for the thermal and the quantum MPM estimations are exactly the same.

We first evaluate the condition, $(\partial p_b/\partial \Gamma_{\rm eff}) = 0$ for $p_b = (1 - m_0)/2 + \sum_\xi \xi \mathcal{M}(\xi) H(u_*)$. After some simple algebra, we obtain

$$m(\Gamma_{\rm eff}^{\rm opt}) \sum_\xi \xi \mathcal{M}(\xi) \exp\left[-\frac{\{a_0 h_* \xi + m(\Gamma_{\rm eff}^{\rm opt})\}^2}{2a^2 h_*^2} \right] = 0. \tag{1.65}$$

Taking into account that $m(\Gamma_{\rm eff}) \neq 0$ is needed for meaningful image restorations, the Nishimori-Wong condition for the quantum MPM estimation is written by

$$\frac{m_0(\beta_s)}{m(\Gamma_{\rm eff}^{\rm opt})} = \frac{a_0}{a^2 h_* \beta_s}. \tag{1.66}$$

As we chose $h_* = \beta_\tau/\beta_s, \beta_\tau = a_0/a^2$, this condition is simply rewritten as $m_0(\beta_s) = m(\Gamma_{\rm eff})$.

Let us summarize the Nishimori-Wong condition for the MPM estimation :

$$\textbf{Thermal :}\ \ T_m^{\rm opt} = T_s \quad \text{(Nishimori and Wong 1999)}$$
$$\textbf{Quantum :}\ \ m_0(\beta_s) = \sum_\xi \mathcal{M}(\xi) \int_{-\infty}^{\infty} \frac{\phi_0 Du}{\sqrt{\phi_0^2 + (\Gamma_{\rm eff}^{\rm opt})^2}}$$

In Fig.1.7, we plot the temperature of the original image T_s-dependence of the optimal temperature $T_m^{\rm opt}$ and the optimal amplitude of the transverse field $\Gamma_{\rm eff}^{\rm opt}$. In the right panel of this figure, the magnetizations $m(T_m^{\rm opt})$ and $m(\Gamma_{\rm eff}^{\rm opt})$ are plotted. The effective amplitude of the transverse field $\Gamma_{\rm eff}$ at which the bit-error rate takes its minimum in Fig. 1.6 is consistent with the $\Gamma_{\rm eff}^{\rm opt}(T_s = 0.9) \simeq 0.66$ as shown in Fig. 1.7 (left).

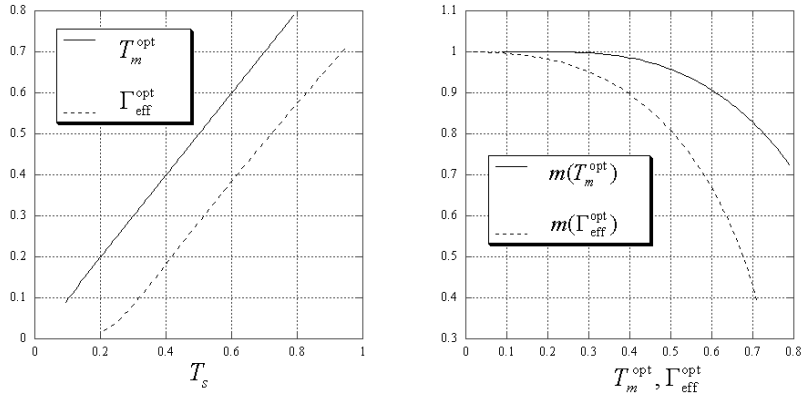

Fig. 1.7. The optimal temperature $T_m^{\rm opt}$ and the optimal transverse field $\Gamma_{\rm eff}^{\rm opt}$ as a function of the temperature T_s of the original image (left), respectively. The optimal temperature for the thermal MPM estimation $T_m^{\rm opt}$ is simply given by $T_m^{\rm opt} = T_s$ (Nishimori temperature). The right panel shows the magnetizations $m(T_m^{\rm opt})$ and $m(\Gamma_{\rm eff}^{\rm opt})$.

From these results, it is shown that the lowest values of the of the bit-error rate for both the thermal and the quantum MPM estimations are exactly the same and the value is given by

$$p_b = \frac{1 - m_0}{2} + \sum_{\xi} \xi \mathcal{M}(\xi) H\left(\frac{a_0 h_* \xi + m_0}{a h_*}\right) \tag{1.67}$$

Therefore, we conclude that it is possible for us to construct the MPM estimation purely induced by the quantum fluctuation (without any thermal fluctuation) and the best possible performance is exactly the same as that of the thermal MPM estimation.

1.4.4 Error-correcting codes

In this subsection, we investigate the performance of the decoding in the so-called Sourlas codes [5], in which uncertainties in the prior are introduced as the quantum transverse field. Although we usually choose the prior in the Sourlas codes as $P(\{\sigma\}) = 2^{-N}$ (the uniform prior), here we use $P(\{\sigma\}) = \prod_i {\rm e}^{-\Gamma \hat{\sigma}_i^x}$. Then, the effective Hamiltonian of the extended Sourlas codes leads to

$$\mathcal{H}_{\rm eff} = -\beta_J \sum_{i1,\cdots,ip} J_{i1\cdots ip}\, \sigma_{i1}^z \sigma_{i2}^z \cdots \sigma_{ip}^z - h \sum_i \tau_i \sigma_i^z - \Gamma \sum_i \sigma_i^x. \tag{1.68}$$

Hereafter, we call this type of error-correcting codes as *Quantum Sourlas codes*. We first derive the Γ-dependence of the bit-error rate for a given p. Then, the channel noise is specified by the next output distribution :

$$P(\{J\},\{\tau\}|\{\xi\}) = \frac{\mathrm{e}^{-\frac{N^{p-1}}{J^2 p!}\sum_{i1,\cdots,ip}\left(J_{i1\cdots ip}-\frac{J_0 p!}{N^{p-1}}\xi_{i1}\cdots\xi_{ip}\right)^2-\frac{1}{2\tau^2}(\tau_i-a_0\xi_i)^2}}{(J^2\pi p!/N^{p-1})^{1/2}\sqrt{2\pi}a} \tag{1.69}$$

For a simplicity, we treat the case in which the original message sequence $\{\xi\}$ is generated by the following uniform distribution $P(\{\xi\}) = 2^{-N}$. Then, the moment of the effective partition function $Z_{\rm eff}$ leads to

$$\begin{aligned} Z_{\rm eff}^n = \exp\Bigg[&\frac{\beta_J}{M}\sum_{i1,\cdots,ip}\sum_{\alpha=1}^{n}\sum_{t=1}^{M} J_{i1\cdots ip}\,\sigma_{i1}^{\alpha}(t)\sigma_{i2}^{\alpha}(t)\cdots\sigma_{ip}^{\alpha}(t) \\ &+ \frac{h}{M}\sum_{i}\sum_{\alpha=1}^{n}\sum_{t=1}^{M}\tau_i\sigma_i^{\alpha}(t) + B\sum_{i}\sum_{t=1}^{M}\sigma_i(t)\sigma_i(t+1)\Bigg] \end{aligned} \tag{1.70}$$

where α and t mean the indexes of the replica number and the Trotter slice, respectively. We set $B \equiv (1/2)\log\coth(\Gamma/M)$ and used the gauge transform : $J_{i1\cdots ip} \to J_{i1\cdots ip}\,\xi_{i1}\cdots\xi_{ip}, \sigma_{ip} \to \xi_{ip}\sigma_{ip}^z$. After averaging $Z_{\rm eff}^n$ over the quenched randomness $[\cdots]_{data}$, namely, over the joint distribution $P(\{J\},\{\tau\},\{\xi\})$, we obtain the following data averaged effective partition function :

$$\begin{aligned} [Z_{\rm eff}^n]_{data} = &\prod_{tt'}\prod_{\alpha\beta}\int_{-\infty}^{\infty}dQ_{\alpha\beta}(t,t')\int_{-\infty}^{\infty}d\lambda_{\alpha\beta}(t,t')\int_{-\infty}^{\infty}dm_{\alpha}(t)\int_{-\infty}^{\infty}d\hat{m}_{\alpha}(t) \\ &\times\exp\left[-Nf(m,\hat{m},\boldsymbol{Q},\boldsymbol{\lambda})\right] \end{aligned} \tag{1.71}$$

with

$$\begin{aligned} f(m,\hat{m},\boldsymbol{Q},\boldsymbol{\lambda}) = &-\frac{\beta_J J_0}{M}\sum_{t,\alpha}m_{\alpha}^p(t) - \frac{h\tau_0}{M}\sum_{t,\alpha}m_{\alpha}(t) \\ &-\frac{(\beta_J J)^2}{4M^2}\sum_{tt',\alpha\beta}Q_{\alpha\beta}^p(t,t') - \frac{(h\tau)^2}{2M^2}\sum_{tt',\alpha\beta}Q_{\alpha\beta}(t,t') \\ &+\frac{1}{M}\sum_{t,\alpha}\hat{m}_{\alpha}(t)m_{\alpha}(t) + \frac{1}{M^2}\sum_{tt',\alpha\beta}\lambda_{\alpha\beta}(t,t')Q_{\alpha\beta}(t,t') \\ &-\frac{1}{M}\sum_{t,\alpha}\hat{m}_{\alpha}(t)\sigma^{\alpha}(t) - \frac{1}{M^2}\sum_{tt',\alpha\beta}\lambda_{\alpha\beta}(t,t')\sigma^{\alpha}(t)\sigma^{\beta}(t') \\ &-B\sum_{t}\sigma(t)\sigma(t+1) \end{aligned} \tag{1.72}$$

where we labeled each Trotter slice by index t. Using the replica symmetric and the static approximations, namely,

$$m_\alpha(t) = m, \quad \hat{m}_\alpha(t) = \hat{m} \tag{1.73}$$

$$Q_{\alpha\beta}(t,t') = \begin{cases} \chi \ (\alpha = \beta) \\ q \ (\alpha \neq \beta) \end{cases}, \quad \lambda_{\alpha\beta}(t,t') = \begin{cases} \lambda_1 \ (\alpha = \beta) \\ \lambda_2 \ (\alpha \neq \beta) \end{cases}, \tag{1.74}$$

we obtain the free energy density f^{RS} :

$$\beta_J f^{RS}(m,\chi,q) = (p-1)\beta_J J_0 m^p + \frac{1}{4}(p-1)(\beta_J J)^2(\chi^p - q^p) - \int_{-\infty}^{\infty} Dw \log \int_{-\infty}^{\infty} Dz \, 2\cosh \Xi \tag{1.75}$$

where we used the saddle point equations with respect to $\hat{m}, \lambda_1, \lambda_2$, namely, $\hat{m} = p\beta_J J_0 m^{p-1} + a_0 h$ and $\lambda_1 = \frac{p}{2}(\beta_J J)^2 \chi^{p-1} + (ah)^2, \lambda_2 = \frac{p}{2}(\beta_J J)^2 q^{p-1} + (ah)^2$. Then, the saddle point equations are derived as follows :

$$m = \int_{-\infty}^{\infty} D\omega \int_{-\infty}^{\infty} Dz \left(\frac{\Phi \sinh \Xi}{\Xi \Omega} \right) \tag{1.76}$$

$$\chi = \int_{-\infty}^{\infty} \frac{D\omega}{\Omega} \int_{-\infty}^{\infty} Dz \left[\left(\frac{\Phi}{\Xi} \right)^2 \cosh \Xi + \Gamma^2 \left(\frac{\sinh \Xi}{\Xi^3} \right) \right] \tag{1.77}$$

$$q = \int_{-\infty}^{\infty} D\omega \left[\int_{-\infty}^{\infty} Dz \left(\frac{\Phi \sinh \Xi}{\Xi \Omega} \right) \right]^2 \tag{1.78}$$

where we defined

$$\Phi = \omega \sqrt{\frac{p}{2}(\beta_J J)^2 q^{p-1} + (ah)^2} + z \sqrt{\frac{p}{2}(\beta_J J)^2 (\chi^{p-1} - q^{p-1})} + p\beta_J J_0 m^{p-1} + a_0 h \tag{1.79}$$

and $\Xi = \sqrt{\Phi^2 + \Gamma^2}$, $\Omega = \int_{-\infty}^{\infty} Dz \cosh \Xi$. The resultant overlap leads to

$$R = \int_{-\infty}^{\infty} D\omega \int_{-\infty}^{\infty} Dz \, \mathrm{sgn}(\Phi) = 1 - 2 \int_{-\infty}^{\infty} Dw \, H(-z_p^*) \tag{1.80}$$

where we defined z_p^* by

$$z_p^* = - \frac{(p\beta_J J_0 m^{p-1} + a_0 h) + w \sqrt{\frac{p}{2}(\beta_J J)^2 q^{p-1} + (ah)^2}}{\sqrt{\frac{p}{2}(\beta_J J)^2 (\chi^{p-1} - q^{p-1})}} \tag{1.81}$$

and the error function $H(x)$ defined as $H(x) = \int_x^{\infty} Dz$. Thus, the bit-error rate for the problem of error-correcting codes is given by $p_b = (1-R)/2 = \int_{-\infty}^{\infty} Dw \, H(-z_p^*)$. where the above bit-error rate p_b depends on Γ through the order parameters χ, q and m.

1.4.5 Analysis for finite p

We first evaluate the performance of the quantum Sourlas codes for the case of finite p by solving the saddle point equations numerically.

Absence of the external field $h = 0$

In Fig. 1.8 (left), we first plot the Γ-dependence of the bit-error rate p_b for the case of $p = 2$ without magnetic field $h = 0$. In this plot, we choose $J = J_0 = 1$ and set $\beta_J = 1$. It must be noted that J_0/J corresponds to the signal to noise ratio (SN ratio). From this figure, we find that the bit error rate gradually approaches to the random guess limit $p_b = 0.5$ as Γ increases. This transition is regarded as a second order phase transition between the ferromagnetic and the paramagnetic phases. We plot the Γ-dependence of the order parameters m, χ and q in the right panel of Fig. 1.8. We should notice that in the classical limit $\Gamma \to 0$, the order parameter χ should takes 1 and both magnetization m and spin glass order parameter q continuously becomes zero at the transition point. Therefore, for the case of $p = 2$, the increase of the quantum fluctuation breaks the error-less state gradually. On

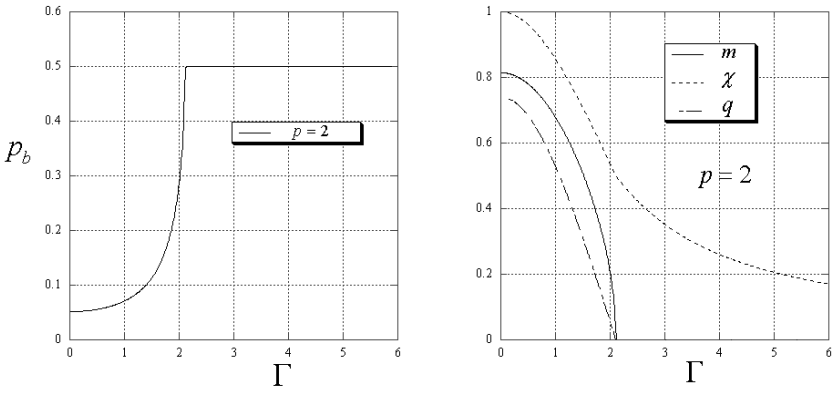

Fig. 1.8. The Γ-dependence of the bit error-rate p_b for the case of $p = 2$ without magnetic field $h = 0$ (left) and and order parameters m, χ and q as a function of Γ (right). We set $\beta_J = 1, J = J_0 = 1$.

the other hand, in Fig. 1.9, we plot the Γ-dependence of the bit-error rate p_b for the case of $p = 3$. In this figure, we find that the bit-error rate suddenly increases to 0.5 at the transition point $\Gamma = \Gamma_c$ and the quality of the message-retrieval becomes the same performance as the random guess. This first order phase transition from the ferromagnetic error-less phase to the paramagnetic random guess phase is observed in the right panel of Fig. 1.9.

We find that the system undergoes the first order phase transition for $p \geq 3$. In Fig. 1.10, we plot the Γ-dependence of the bit-error rate for $p =$

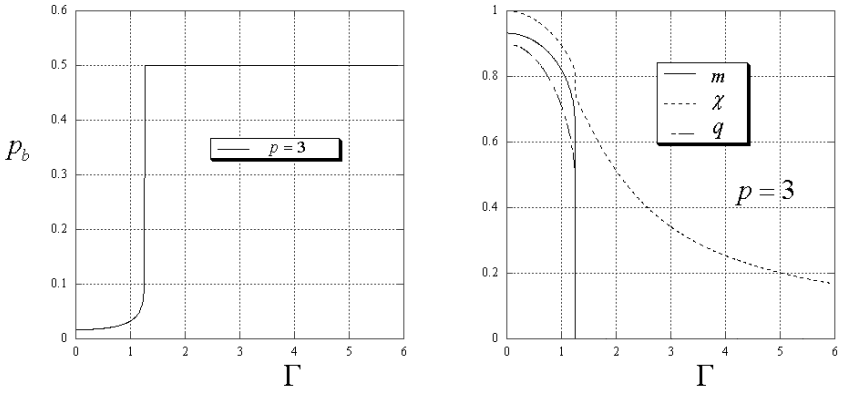

Fig. 1.9. The Γ-dependence of the bit error-rate p_b for the case of $p = 3$ without magnetic field $h = 0$ (left) and and order parameters m, χ and q as a function of Γ (right). We set $\beta_J = 1, J = J_0 = 1$.

$2, 3, \cdots, 6$ and $p = 12$. From this figure, we find that the transition for $p \geq 3$ is first order and the bit-error rate changes its state from the ferro-magnetic almost perfect information retrieval phase to the paramagnetic random guess phase at $\Gamma = \Gamma_c$. The tolerance to the quantum fluctuation increases as the number of degree p of the interaction increases.

Presence of the external field $h \neq 0$

We next consider the case of $h \neq 0$. This means that we send not only the parity check $\{J_{i1\cdots ip}\}$ but also bit sequence $\{\xi\}$ itself. We plot the bit-error rate as a function of Γ in Fig. 1.11. From this figure, we find that the bit-error rate goes to some finite value which is below the random guess limit gradually. The right panel of this figure tells us that in this case there is no sharp phase transition induced by the quantum fluctuation. In Fig. 1.12, we plot the bit-error rate and corresponding order parameters as a function of Γ. This figure tells us that the bit-error rate suddenly increases at some critical length of the transverse field Γ_c. As we add the external field h, this is not a ferro-para magnetic phase transition, however, there exist two stable states, namely good retrieval phase and poor retrieval phase. In Fig. 1.13, we plot the Γ-dependence of the bit-error rate for $p = 3, \cdots, 6$ and $p = 12$ (left) and for $p = 6$ and $\beta_J = 0.2, \cdots, 12$ (right). From this right panel, interesting properties are observed. For small Γ, the bit-error rate becomes small as we increases p. On the other hand, for large Γ, the bit error rate becomes large

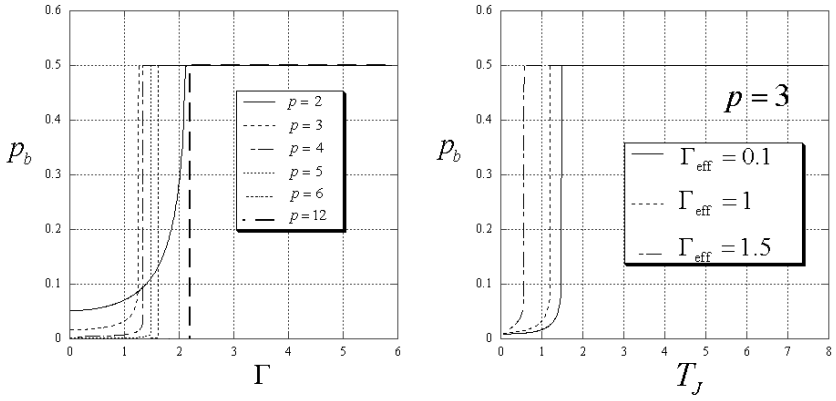

Fig. 1.10. The Γ-dependence of the bit error-rate p_b for $p = 2, \cdots, 6$ and $p = 12$ without magnetic field $h = 0$ (left). We set $\beta_J = 1, J = J_0 = 1$. The right panel shows the $T_J = \beta_J^{-1}$ dependence of the bit-error rate for keeping the ratio : $\Gamma/\beta_J \equiv \Gamma_{\rm eff}$ to the values $\Gamma_{\rm eff} = 0.1, 1$ and 1.5.

as p increases. Moreover, the bit-error rate for $p = 6$ takes its maximum at some finite value of Γ.

1.4.6 Phase diagrams for $p \to \infty$ and replica symmetry breaking

In this subsection, we investigate properties of the quantum Sourlas codes in the limit of $p \to \infty$. In this limit, we easily obtain several phase boundaries analytically and draw the phase diagrams.

First of all, we consider the simplest case, namely, the case of $J_0 = 0, h = 0$. For this choice of parameters, the ferromagnetic phase does not appear and possible phases are paramagnetic phase and spin glass phase. The free energy density we evaluate is now rewritten by

$$f^{RS} = -\frac{1}{4}(p-1)\beta_J J^2(q^p - \chi^p) - T_J \int_{-\infty}^{\infty} Dw \log \int_{-\infty}^{\infty} Dz 2\cosh \beta_J \sqrt{\phi_0^2 + \Gamma_{\rm eff}^2} \tag{1.82}$$

with $\phi_0 = w\sqrt{pJ^2 q^{p-1}/2} + z\sqrt{pJ^2(\chi^{p-1} - q^{p-1})/2}$, where we defined $\Gamma_{\rm eff} = \Gamma/\beta_J$. In the paramagnetic phase, there is no spin glass ordering, namely, $q = 0$. Thus, the free energy density in the paramagnetic phase leads to

$$f_{para}^{RS} = \frac{J^2\beta_J}{4}(p-1)\chi^p - T_J \log \int_{-\infty}^{\infty} Dz 2\cosh \beta_J \sqrt{\Gamma_{\rm eff}^2 + \frac{p}{2}J^2\chi^{p-1}z^2}. \tag{1.83}$$

The saddle point equation with respect to χ is given by

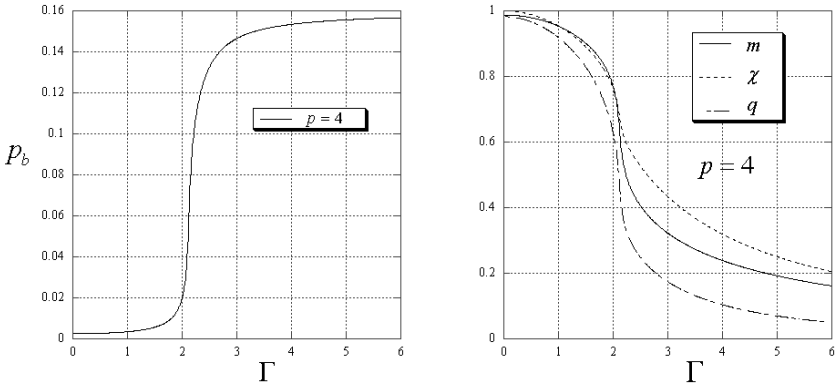

Fig. 1.11. The Γ-dependence of the bit error-rate p_b for the case of $p = 4$ with magnetic field $h = 1$ (left) and and order parameters m, χ and q as a function of Γ (right). We set $\beta_J = 1, J = J_0 = 1$ and $a_0 = a = 1$.

$$\chi = \frac{\int_{-\infty}^{\infty} Dz \left\{ \left(\frac{\phi_{00}}{\Gamma_{\rm eff}^2 + \phi_{00}} \right) \cosh \beta_J \sqrt{\Gamma_{\rm eff}^2 + \phi_{00}} + \Gamma_{\rm eff}^2 T_J \frac{\sinh \beta_J \sqrt{\Gamma_{\rm eff}^2 + \phi_{00}}}{\sqrt{\Gamma_{\rm eff}^2 + \phi_{00}}^3} \right\}}{\int_{-\infty}^{\infty} Dz \cosh \beta_J \sqrt{\Gamma_{\rm eff}^2 + \phi_{00}^2}} \tag{1.84}$$

with $\phi_{00} = pJ^2 \chi^{p-1} z^2/2$. In the limit of $p \to \infty$, there are two possible solutions of χ, that, is $\chi^p = 1$ and $\chi^p = 0$. The former is explicitly given from (1.84) as $\chi \simeq 1 - 4\Gamma_{\rm eff}^2 T_J^2/p^2 J$. Then, we obtain the free energy density for this solution as $f_I = -J^2/4T_J - T_J \log 2$ by substituting this χ into (1.83) and evaluating the integral with respect to z at the saddle point in the limit of $p \to \infty$. Let us call this phase as *PI*. The later solution is explicitly evaluated as $\chi = (T_J/\Gamma_{\rm eff}) \tanh(\Gamma_{\rm eff}/T_J)$ $(< 1, \text{thus}, \chi^p = 0)$ and corresponding free energy density leads to $f_{II} = -T_J \log 2 - T_J \log \cosh(\Gamma_{\rm eff}/T_J)$. We call this phase as *PII*.

Here we should not overlook the entropy in *PI*, namely, $S = -(\partial f_I/\partial T) = -J^2/4T_J^2 + \log 2$. Obviously, S becomes negative for $T < (J/2\sqrt{\log 2})^{-1}$ and in this region, the replica symmetry of the order parameters might be broken. Therefore, in this low temperature region, we should construct the replica symmetry breaking (RSB) solution. To obtain the RSB solution, we break the symmetry of the matrices $\boldsymbol{q}$ and $\boldsymbol{\lambda}$ as

$$q_{l\delta,l'\delta'} = \begin{cases} q_0 \ (l = l') \\ q_1 \ (l \neq l') \end{cases}, \quad \lambda_{l\delta,l'\delta'} = \begin{cases} \hat{\lambda}_0 \ (l = l') \\ \hat{\lambda}_1 \ (l \neq l') \end{cases} \tag{1.85}$$

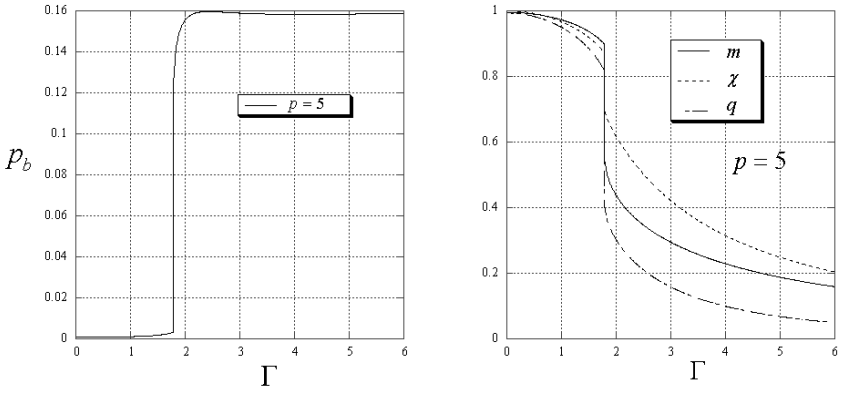

Fig. 1.12. The Γ-dependence of the bit error-rate p_b for the case of $p = 5$ with magnetic field $h = 1$ (left) and and order parameters m, χ and q as a function of Γ (right). We set $\beta_J = 1, J = J_0 = 1$ and $a_0 = a = 1$.

for $l = 1, \cdots, n/x$, $\delta = 1, \cdots, x$. Then, we obtain the free energy density for one step RSB solution as

$$\begin{aligned} f^{1RSB} &= (p-1)J_0 m^p + \frac{\beta_J J^2}{4}[x q_1^p + (1-x)q_0^p] + \frac{\beta_J J^2}{4}(p-1)\chi^p \\ &- \frac{\beta_J}{2}[x q_1 \hat{\lambda}_1 + (1-x) q_0 \hat{\lambda}_0] \\ &- \frac{T_J}{x}\int_{-\infty}^{\infty} Dw \log \int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh \beta_J \sqrt{\hat{\phi}^2 + \Gamma_{\rm eff}^2} \right)^x \quad (1.86) \end{aligned}$$

with $\hat{\phi} = w\sqrt{\hat{\lambda}_1} + z\sqrt{\hat{\lambda}_0 - \hat{\lambda}_1} + y\sqrt{pJ^2\chi^{p-1}/2 - \hat{\lambda}_0} + p\beta_J J_0 m^{p-1} + a_0 h$. By taking $(\partial f^{1RSB}/\partial q_0) = (\partial f^{1RSB}/\partial q_1) = 0$, we obtain $\hat{\lambda}_1 = pJ^2 q_1^{p-1}/2, \hat{\lambda}_0 = pJ^2 q_0^{p-1}/2$.

Here we set the parameters J_0, h again to $J_0 = h = 0$. At low temperature, we naturally assume $q_1 < 0$ $(\hat{\lambda}_1 = 0), q_0 = 1$ $(\hat{\lambda}_0 = pJ^2/2)$ and $\chi = 1$. Substituting these conditions into (1.86) and evaluating the integral with respect to y at the saddle point in the limit of $p \to \infty$, we obtain the free energy density in this phase, which will be referred to as SGI, as $f_{SGI} = -\beta_J J^2 x/4 - \log 2/(\beta_J x)$. Substituting the solution of $(\partial f_{SGI}/\partial x) = 0$, namely, $x = 2\sqrt{\log 2}/(JT_J)$ into f_{SG}, we obtain the free energy density which specifies SGI as $f_{SGI} = -J\sqrt{\log 2}$.

Let us summarize :

$$PI \text{ (para)} : f_I = -\frac{J^2}{4T_J} - T_J \log 2 \quad (\chi = 1, q = 0)$$

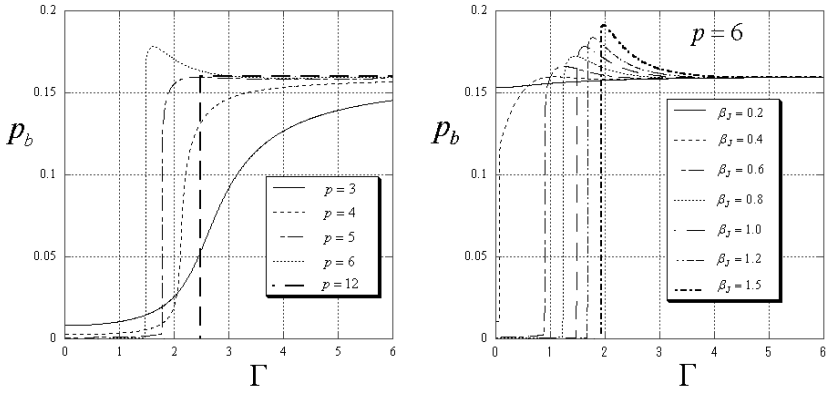

Fig. 1.13. The Γ-dependence of the bit-error rate for $p = 3, \cdots, 6$ and 12 (left). For $p = 6$, the Γ-dependences of the bit-error rate for $\beta_J = 0.2, \cdots, 1.2$ are shown in the right panel.

$$PII \text{ (para)} : f_{II} = -T_J \log 2 - T_J \log \cosh \left(\frac{\Gamma_{\text{eff}}}{T_J} \right) \quad (\chi^p = q = 0)$$

$$SGI \text{ (spin glass)} : f_{SGI} = -J\sqrt{\log 2} \quad (\chi = q = 1)$$

We illustrate the phase diagram in Fig.1.14 (left).

As the phase transitions between arbitrary two phases among these three (*PI,PII,SGI*) are all first order, each phase boundary is obtained by balancing of the free energy density. Namely, $\Gamma_{\text{eff}} = T_J \cosh^{-1}(\mathrm{e}^{J^2/4T_J^2})$ $(T_J > T_c)$ for *PI-PII*, $\Gamma_{\text{eff}} = T_J \cosh^{-1}(\mathrm{e}^{J\sqrt{\log 2}/T_J}/2)$ $(T_J < T_c)$ for *PII-SGI* and $T_J = J/2\sqrt{\log 2} = T_c$ for *SGI-PI*.

We next consider the case of $J_0 \neq 0$. This case is much more important in the context of error-correcting codes. For the case of absence of the external field $h = 0$, the phase transition between the error-less phase and the random guess phase is specified as the ferro-paramagnetic (or spin glass) phase transition. From reasons we mentioned above, our main purpose here is to determine the transition point $(J_0/J)_c$ below which the ferromagnetic phase is stable. The critical SN ratio $(J_0/J)_c$ is important because as we mentioned before, the error-less decoding is possible when the channel capacity C and the transmission rate R satisfy the inequality $R \leq C$. The channel capacity for the Gaussian channel we are dealing with is given by $C = (1/2)\log_2(1 + J_0^2/J^2)$ with $J_0 = J_0 p!/N^{p-1}, J = J^2 p!/2N^{p-1}$, that is, $C \simeq J_0^2 p!/(J^2 N^{p-1} \log 2)$ in the limit of $N \to \infty$ for a given p. On the other hand, the transmission rate R is given as $R = N/N_B = N/{}_N C_p \simeq p!/N^{p-1}$.

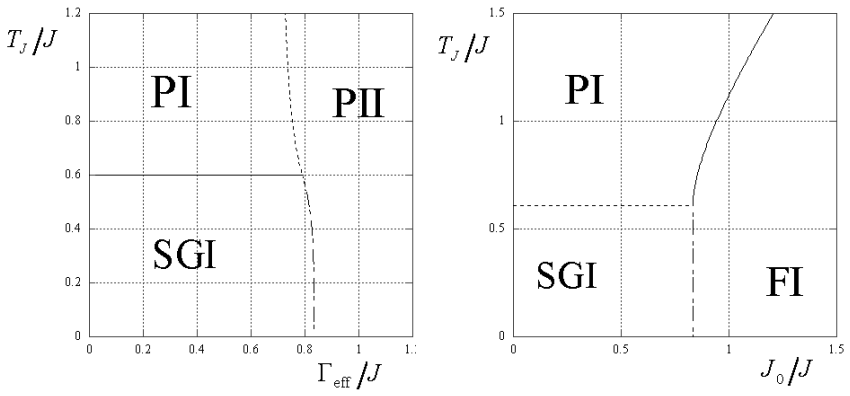

Fig. 1.14. The phase diagrams in the limit of $p \to \infty$. In the case of $J_0 = 0$, there exist three phases, namely, *PI*,*PII* and *SGI*. Below the critical point $(T_J/J)_c = (1/2J\sqrt{\log 2})$, the replica symmetry is broken (the left panel). The right panel shows the critical SN ratio $(J_0/J)_c$, above which (labeled *FI* in the panel) decoding without errors is achieved, is given by $\sqrt{\log 2}$.

Therefore, the error-less decoding is possible when the following inequality :

$$\frac{R}{C} = \left(\frac{J}{J_0}\right)^2 \log 2 \leq 1 \tag{1.87}$$

holds and the question now arises, namely, it is important to ask whether the above inequality is satisfied or not at the critical point $(J/J_0)_c$. In following, we make this point clear.

We start from the saddle point equations which are derived from the free energy density of the one step RSB (1.86). These equations are given explicitly as

$$\begin{aligned}
m &= \int_{-\infty}^{\infty} Dw \frac{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J\hat{\Xi}\right)^{x-1} \int_{-\infty}^{\infty} Dy \left(\frac{\hat{\phi}}{\hat{\Xi}}\right) 2\sinh\beta_J\hat{\Xi}}{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J\hat{\Xi}\right)^{x}} \\
q_0 &= \int_{-\infty}^{\infty} Dw \frac{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J\hat{\Xi}\right)^{x-2} \left(\int_{-\infty}^{\infty} Dy \left(\frac{\hat{\phi}}{\hat{\Xi}}\right) 2\sinh\beta_J\hat{\Xi}\right)^2}{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J\hat{\Xi}\right)^{x}} \\
q_1 &= \int_{-\infty}^{\infty} Dw \left\{ \frac{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy \left(\frac{\hat{\phi}}{\hat{\Xi}}\right) 2\sinh\beta_J\hat{\Xi}\right)^{x}}{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J\hat{\Xi}\right)^{x}} \right\}^2
\end{aligned} \tag{1.88}$$

$$\chi = \int_{-\infty}^{\infty} Dw \frac{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J \hat{\Xi} \right)^{x-1} \int_{-\infty}^{\infty} Dy \left(\frac{\hat{\phi}}{\hat{\Xi}} \right)^2 2\sinh\beta_J \hat{\Xi}}{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J \hat{\Xi} \right)^{x}}$$
$$+ \Gamma_{\rm eff}^2 T_J \int_{-\infty}^{\infty} Dw \frac{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J \hat{\Xi} \right)^{x-1} \int_{-\infty}^{\infty} Dy \left(\frac{2\sinh\beta_J \hat{\Xi}}{\hat{\Xi}^3} \right)}{\int_{-\infty}^{\infty} Dz \left(\int_{-\infty}^{\infty} Dy 2\cosh\beta_J \hat{\Xi} \right)^{x}} \tag{1.89}$$

with

$$\hat{\phi} = Jw\sqrt{\frac{p}{2}q_1^{p-1}} + Jz\sqrt{\frac{p}{2}(q_0^{p-1} - q_1^{p-1})} + Jy\sqrt{\frac{p}{2}(\chi^{p-1} - q_0^{p-1})} + pJ_0 m^{p-1} \tag{1.90}$$

and $\hat{\Xi} = \sqrt{\hat{\phi}^2 + \Gamma_{\rm eff}^2}$. When the number of product p of the estimate of the original bits is extremely large and $J/J_0, m$ is positive, $\hat{\phi} = pJ_0 m^{p-1}$ and the solutions of the above saddle point equations lead to $m = q_0 = q_1 = 1$ and $\chi = 1$. Thus, the system is in the ferromagnetic phase and the replica symmetry is not broken ($q_0 = q_1$). Substituting the replica symmetric solution $m = q = 1$ into (1.75) and evaluating the integral with respect to w at the saddle point in the limit of $p \to \infty$, we obtain the free energy density in this phase (let us call *FI*) as $f_{FI} = -J_0$. We should notice that this free energy density does not depend on the effective amplitude of the transverse field $\Gamma_{\rm eff}$ at all. From the argument of $J_0 = 0$ case, the phase specified $\chi = 1, T_J < T_c = (2\sqrt{\log 2})$ is spin glass phase. Therefore, the condition (1.87) is satisfied and the ferromagnetic error-less phase exists for $(J_0/J) \geq (J_0/J)_c = \sqrt{\log 2}$, where (J_0/J) is determined by balancing of the free energy densities $f_{FI} = f_{SGI}$. As the result, we conclude that the error-less decoding is achieved if the SN ratio (J_0/J) is greater than the critical value $(J_0/J)_c = \sqrt{\log 2}$ and the condition is independent of $\Gamma_{\rm eff}$. To put it into another word, the Shannon's bound is not violated by the quantum uncertainties in the prior distribution in the limit of $p \to \infty$.

The details of the analysis, including the numerical RSB solutions for finite p will be reported in the conference and in forth coming article [28].

1.5 Quantum Markov chain Monte Carlo simulation

In the previous section, we investigated the performance of the MAP and the MPM estimations for the problems of image restoration and error-correcting codes by using analysis of the mean-field infinite range model. In Sourlas codes, the infinite range model is naturally accepted because we do not have to consider any structure in the bit sequence $\{\xi\}$, and in that sense, the range of interactions in the parity check $\{\xi_{i1} \cdots \xi_{ip}\}$ is infinite.

On the other hand, in image restoration, there should exist some geometrical structures in each pair in the sequence of the original image $\{\xi\}$. Then, we should introduce appropriate two dimensional lattice on which each pixel is located. Therefore, in this section, we carry out computer simulations for the two dimensional model system to investigate the qualities of the MAP and the MPM image restorations quantitatively.

1.5.1 Quantum Markov chain Monte Carlo method

Let us remind of readers that our effective Hamiltonian for image restoration is described by $\mathcal{H}_{\rm eff} = -\beta_m \sum_{<ij>} \sigma_i^z \sigma_j^z - h \sum_i \tau_i \sigma_i^z - \Gamma \sum_i \sigma_i^x$. In this section, we suppose that each pixel σ_i^z is located on the two dimensional square lattice. To evaluate the expectation value of arbitrary quantity A in the quantum spin system

$$\langle A \rangle = \frac{{\rm tr}_{\{\sigma\}}\, A\, {\rm e}^{-\beta \mathcal{H}_{\rm eff}}}{{\rm tr}_{\{\sigma\}}\, {\rm e}^{-\beta \mathcal{H}_{\rm eff}}}, \tag{1.91}$$

we use the following ST formula [24] to carry out the above trace in practice as

$$\exp(-\beta \mathcal{H}_{\rm eff}) = \lim_{M \to \infty} \left({\rm e}^{\frac{\mathcal{A}}{M}} {\rm e}^{\frac{\mathcal{B}}{M}} \right) \tag{1.92}$$

where we defined

$$\mathcal{A} = \beta(\beta_m \sum_{<ij>} \sigma_i^z \sigma_j^z + h \sum_i \tau_i \sigma_i^z) = -\beta \mathcal{H}_{\rm eff}^{classical}, \;\; \mathcal{B} = \beta \Gamma \sum_i \sigma_i^x. \tag{1.93}$$

We should keep in mind that these two terms $\mathcal{A}$ and $\mathcal{B}$ are easily diagonalized.

Then, by inserting the complete set : $\sum_{\{\sigma_{jk}\}} |\{\sigma_{jk}\}\rangle\langle\{\sigma_{jk}\}| = 1$, the partition function Z_M for a fixed Trotter size M leads to

$$\begin{aligned} Z_M = {\rm tr}_{\{\sigma\}} \left({\rm e}^{\frac{\mathcal{A}}{M}} {\rm e}^{\frac{\mathcal{B}}{M}} \right) = \sum_{\{\sigma_{jk}=\pm 1\}} & \langle\{\sigma_{j1}\}|{\rm e}^{\frac{\mathcal{A}}{M}}|\{\sigma_{j1}^{'}\}\rangle\langle\{\sigma_{j1}^{'}\}|{\rm e}^{\frac{\mathcal{B}}{M}}|\{\sigma_{j2}\} \times \cdots \\ & \times \cdots \times \langle\{\sigma_{jM}\}|{\rm e}^{\frac{\mathcal{A}}{M}}|\{\sigma_{jM}^{'}\}\rangle\langle\{\sigma_{jM}^{'}\}|{\rm e}^{\frac{\mathcal{B}}{M}}|\{\sigma_{j1}\}\rangle \end{aligned} \tag{1.94}$$

where $|\{\sigma_{jk}\}\rangle$ is M-th product of eigenvectors $\{\sigma\}$ and is explicitly given by $|\{\sigma_{jk}\}\rangle = |\sigma_{j1}\rangle \otimes |\sigma_{j2}\rangle \otimes \cdots \otimes |\sigma_{jM}\rangle$.

By taking the limit of $M \to \infty$, we obtain the *effective partition function* $Z_{\rm eff}$ of the quantum spin system with $B = (1/2)\log\coth(\beta\Gamma/M)$ as follows.

$$\begin{aligned} & Z_{\rm eff} \equiv \lim_{M \to \infty} Z_M \\ & = \lim_{M \to \infty} (a_M)^N \sum_{\{\sigma_{jk}\}=\pm 1} {\rm e}^{\frac{\beta\beta_m}{M} \sum_{ij,k} \sigma_i^k \sigma_j^k + \frac{\beta h}{M} \sum_{i,k} \tau_i \sigma_i^k + B \sum_{i,k} \sigma_i^k \sigma_i^{k+1}} \\ & = \lim_{M \to \infty} (a_M)^N \end{aligned}$$

$$\times \sum_{\{\sigma_{jk}=\pm 1\}} \exp\left[\beta_{\rm eff}\left\{\beta_m \sum_{ij,k}\sigma_i^k\sigma_j^k + h\sum_{i,k}\tau_i\sigma_i^k + B_M\sum_{i,k}\sigma_i^k\sigma_i^{k+1}\right\}\right] \tag{1.95}$$

where we defined a_M and B_M as $a_M \equiv \{(1/2)\sinh(2\beta_{\rm eff}\Gamma)\}^{1/2}$, $B_M \equiv (1/2\beta_{\rm eff})\log\coth(\beta_{\rm eff}\Gamma)$ and introduced the following *effective inverse temperature* : $\beta_{\rm eff} = \beta/M$. Thus, this is the partition function of a $(d+1)$-dimensional classical system at the effective temperature $T_{\rm eff} = \beta_{\rm eff}^{-1}$.

Let us think about the limit of $\Gamma \to 0$ in this expression. Then, the coupling constant of the last term appearing in the argument of the exponential becomes strong. As the result, copies of the original system, which are described by the $\mathcal{H}_{\rm eff}^{classical}$ and located in the Trotter direction labeled by k, have almost the same spin configurations. Thus, the partition function is now reduced to that of the classical system at temperature $T = \beta^{-1}$.

We should not overlook that when we describe the same quantum system at $T = 0$ of the effective Hamiltonian $\mathcal{H}_{\rm eff}^{Quantum}$ by analysis of Schrödinger equation : $i\hbar(\partial|\psi(t)\rangle/\partial t) = \mathcal{H}(t)|\psi(t)\rangle$ for the time dependent Hamiltonian : $\mathcal{H}(t) = -\beta_m\sum_{<ij>}\sigma_i^z\sigma_j^z - h\sum_i\tau_i\sigma_i^z - \Gamma(t)\sum_i\sigma_i^x$, the inverse temperature β does not appear in the above expression. Therefore, we can not use β in the quantum Monte Carlo method to simulate the quantum system at $T = 0$.

To realize the equilibrium state at the ground state $T = 0$ for a finite amplitude of the quantum fluctuation $\Gamma \neq 0$, we take the limit $\beta \to \infty, M \to \infty$ keeping the effective inverse temperature $\beta_{\rm eff} = \mathcal{O}(1)$. Namely, effective parameters to simulate the pure quantum system by the quantum Monte Carlo method are $\beta_{\rm eff}$ and M, instead of β and M. This choice is quite essential especially in the procedure of quantum annealing [16] because the quantum annealing searches the globally minimum energy states by using only the quantum fluctuation without any thermal fluctuation. Therefore, if we set the effective inverse temperature $\beta_{\rm eff}$ as of order 1 object in the limit of $M \to \infty$ (we can take into account the quantum effect correctly in this limit) and $\beta \to \infty$ (the thermal fluctuation is completely suppressed in this limit), we simulate the quantum spin system at the ground state $T = 0$.

1.5.2 Quantum annealing and simulated annealing

According to the argument in the previous subsection, we construct the quantum annealing algorithm to obtain the globally minimum energy states of our effective Hamiltonian $\mathcal{H}_{\rm eff}^{classical}$. To realize the algorithm, we control the amplitude of the transverse field as

$$\textbf{Quantum Annealing (QA)} : \Gamma \to 0 \quad \text{for } \beta_{\rm eff} = 1,\ M \to \infty$$

We should notice that the simulated annealing (thermal annealing) is achieved by controlling the parameter β as

Simulated Annealing (SA) : $\beta \to \infty$ for finite M and $\Gamma = 0$.

As we mentioned, the scheduling of $T(t)$ and $\Gamma(t)$ might be essential in the simulated annealing and the quantum annealing. Although we know the optimal temperature scheduling $T(t) \sim (\log t)^{-1}$, however, we do not yet obtain any mathematically rigorous arguments for $\Gamma(t)$ as in the simulated annealing. Therefore, in this section, we use the same scheduling for $\Gamma(t)$ as that of the simulated annealing, namely, $T(t) = \Gamma(t)$. The justification of identification of $\Gamma(t)$ and $T(t)$ comes from the results we obtained in the previous section, that is, the shape of the bit-error rate at $T = 0$ as a function of Γ is almost same as the bit-error rate for the thermal one. Thus, we assume that Γ and T might have the same kind of role to generate the equilibrium states for a given Γ and T. However, the mathematical arguments on the scheduling of Γ are quite important and should be made clear in near future.

1.5.3 Application to image restoration

We investigate the MAP and MPM estimations by the quantum Monte Carlo method and the quantum annealing for the two dimensional pictures which are generated by the Gibbs distribution : $P(\{\xi\}) = \mathrm{e}^{\beta_s \sum_{<ij>} \xi_i \xi_j} / Z(\beta_s)$. It must be noted that in the above sum $\sum_{<ij>}(\cdots)$ should be carried out for the nearest neighboring pixels located on the two dimensional square lattice. A typical snapshot from this distribution is shown in Fig. 1.16.

Thermal MPM estimation versus quantum MPM estimation

Before we investigate the performance of the simulated annealing and the quantum annealing, as a simple check for our simulations, we demonstrate the thermal MPM estimation for the degraded image with $p_\tau = 0.1$ of the original image generated at $T_s = 2.15$ by using the thermal and the quantum Markov chain Monte Carlo methods. We show the result of the T_m-dependence of the bit-error rate in Fig.1.15. We carried out 30-independent runs for system size 100×100. We set $h/\beta_m = T_s \beta_\tau = (T_s/2) \log(1 - p_\tau/p_\tau)$. From this figure, we find that the best performance is achieved around the temperature $T_m = T_s = 2.15$. In Fig.1.16, we show the original, the degraded and restored images. From this figure, we found that the restored image at relatively low temperature $T_m = 0.6$ is pained in even for the local structure of the original images. On the other hand, at the optimal temperature $T_m = 2.15$, the local structures of the original image are also restored.

We next investigate the quantum MPM estimation. In Fig.1.15, we plot the bit-error rate for the quantum MPM estimation of the original image generated by the Gibbs distribution for the two dimensional ferromagnetic Ising model. We control the effective transverse field Γ_{eff} on condition that the inverse temperature β is setting to $\beta = \beta_{\mathrm{eff}} M$, namely, the effective inverse temperature $\beta_{\mathrm{eff}} = 1$. The hyperparameter $\beta_m^{-1} = T_m$ and h are fixed to their

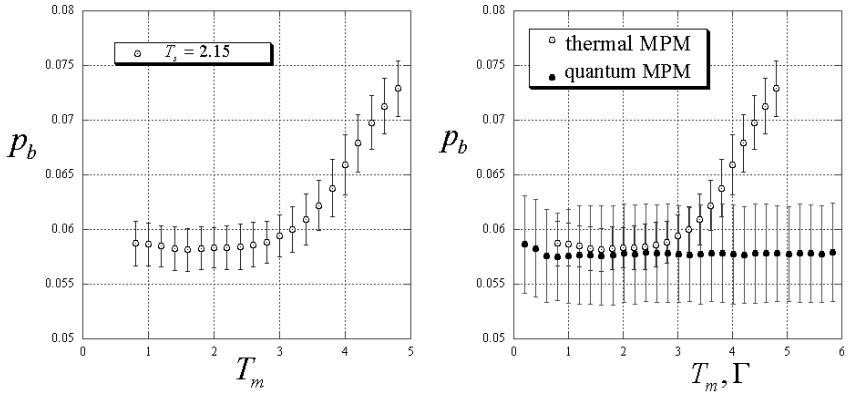

Fig. 1.15. The bit-error rate p_b for the thermal MPM estimation as a function of the temperature T_m (left). The plots were obtained from 30-independent runs for the system size 100×100. We set the temperature of the original image $T_s = 2.15$ and the noise rate $p_\tau = 0.1$. The right panel shows the bit-error rate for the quantum MPM estimation for the system size 50×50, and the Trotter number $M = 200$ for the same noise level $p_\tau = 0.1$ as the left panel. The error-bars are obtained from 50-independent runs.

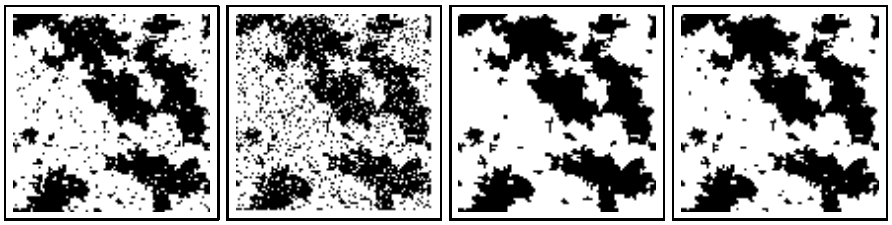

Fig. 1.16. From the left to the right, the original, the degraded ($p_\tau = 1$), and the restored at $T_m = 0.6$ and $T_m = T_s = 2.15$ pictures are displayed.

optimal values $T_m = T_s = 2.15$ and $h = \beta_\tau = (1/2)\log(1 - p_\tau/p_\tau)$. To draw this figure, we carry out 50-independent runs for the system size 50×50 for the Trotter size $M = 200$. The Monte Carlo Step (MCS) needed to obtain the equilibrium state is chosen as $t^{'} = Mt$, where $t = 10^5$ is the MCS for the thermal MPM estimation. One Monte Carlo step in calculation the quantum MPM estimate takes M times evaluations of spin flips than the calculation of the thermal MPM estimate. Thus, we provide a reasonable definition of the time $t^{'}$ of which the quantity is plotted and compared as a function as $t^{'} = t$ (thermal) and $t^{'} = Mt$ (quantum).

From this figure, we find that the lowest values of the bit-error rate for the quantum and the thermal MPM estimations are almost the same value as our analysis of the mean-field infinite range model predicted, however, the Γ-dependence of the bit-error rate is almost flat. We display several typical examples of restored images by the thermal and quantum MPM estimations in Fig. 1.17 From this figure, we find that the performance of the quantum

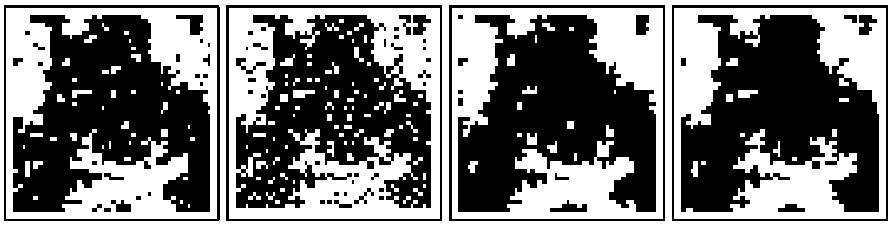

Fig. 1.17. From the left to the right, 50×50 original image generated at $T_s = 2.15$, degraded images ($p_\tau = 0.1$), and restored image by the thermal MPM estimation, and the restored image by the quantum MPM estimation. Each bit-error rate is $p_b = 0.06120$ for the thermal MPM at $T_m = T_s = 2.15$ and $p_b = 0.06040$ for the quantum MPM estimation with $\Gamma = 0.8$ (at the nearest point form the solution of $m_0 = m(\Gamma)$), respectively.

MPM estimation is slightly superior to the thermal MPM.

Simulated annealing versus quantum annealing

In last part of this section, we investigate how effectively the quantum tunneling process possibly leads to the global minimum of the effective Hamiltonian for the image restoration problem in comparison to temperature-driven process used in the simulated annealing. It is important for us to bear in mind that the observables we should check in the problem of image restoration are not only the energy on time E but also the bit-error rate p_b. As we mentioned, the globally minimum energy state of the classical Hamiltonian does not always minimize the bit-error rate. Therefore, from the view point of image restoration, the dynamics of the bit-error rate is also relevant quantity, although, to evaluate the performance of the annealing procedure, the energy on time is much more important measure. In this article, we investigate both of these two measures.

In our simulations discussed below, we choose the temperature and the amplitude of transverse scheduling as $\Gamma(t) = T(t) = 3/\sqrt{t}$ according to Kadowaki and Nishimori [16]. To suppress the thermal and the quantum fluctuation at the final stage of the annealing procedure, we set $\Gamma = T = 0$ in last 10% of the MCS.

In Fig. 1.18, we plot the time development of the bit-error rate and the energy on time, namely, $E_t = -\beta_m \sum_{<ij>} \sigma_i^z \sigma_j^z - h \sum_i \tau_i \sigma_i^z$, where we defined $\sigma_i^z = (1/M) \sum_k \sigma_i^k$ for the quantum annealing. As the MCS $t^{'}$ for the quantum annealing is defined by $t^{'} = Mt$ for the MCS, where t is the MCS for the SA, we should not overlook that the initial behavior of the first M-th MCS in the quantum annealing is not shown in this figure. We carried out this simulation for system size 50×50 with Trotter size $M = 200$. The noise rate is $p_\tau = 0.1$. We set $\beta_m^{-1} = T_s = 2.15$ and $h = (1/2)\log(1 - p_\tau/p_\tau) = 1.1$. From this figure, we find that the mean value of the bit-error rate calculated by the quantum annealing is smaller than that of the simulated annealing. However, the energy on time of the simulated annealing is slightly lower than that of the quantum annealing. Although this result is not enough to decide which annealing is superior, the simulated annealing with temperature scheduling $T(t) = 3/\sqrt{t}$ seems to be much more effective than the quantum annealing with the same scheduling of the amplitude of the transverse field for finding the minimum energy state. Of course, we should check more carefully to choose the optimal or much more effective scheduling of Γ. This might be one of the important future problems. In Fig.1.19, we display the resultant restored images by the sim-

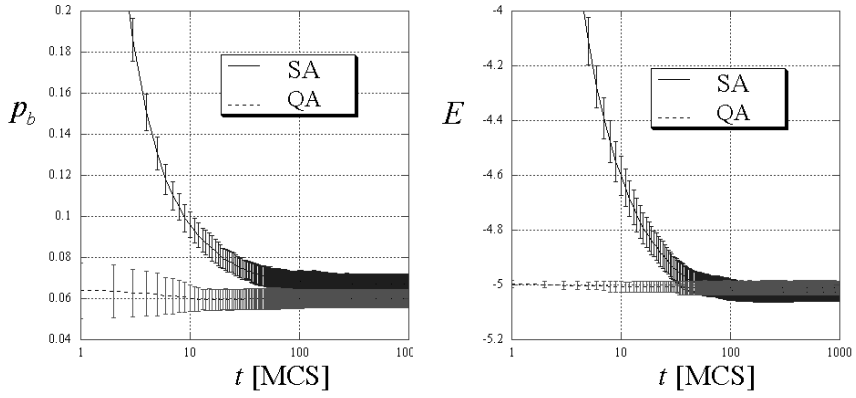

Fig. 1.18. The time dependence of the bit-error rate for the simulated annealing (SA) and the quantum annealing (QA). The MCS $t^{'}$ for the quantum annealing is defined by $t^{'} = Mt$ for the MCS, where t is the MCS for the SA. The right panel indicates the dynamical process of the energy function by the SA and the QA. We carried out this simulation for system size 50×50 with the Trotter size $M = 200$. The noise rate is $p_\tau = 0.1$. The error-bars are calculated by 50-independent runs.

ulated annealing and the quantum annealing. For this typical example, the performance of the quantum annealing restoration measured by the bit-error rate is better than that of the simulated annealing. The difference of the correct pixels is estimated as $\Delta n = 50 \times 50 \times \Delta p_b = 2500 \times 0.0084 = 21$ (pixels), where $\Delta p_b = p_b(\mathrm{SA}) - p_b(\mathrm{QA})$. From reasons we mentioned above, the MAP estimate obtained by the quantum annealing is not a correct MAP estimate, however, the quality of the restoration is really fine.

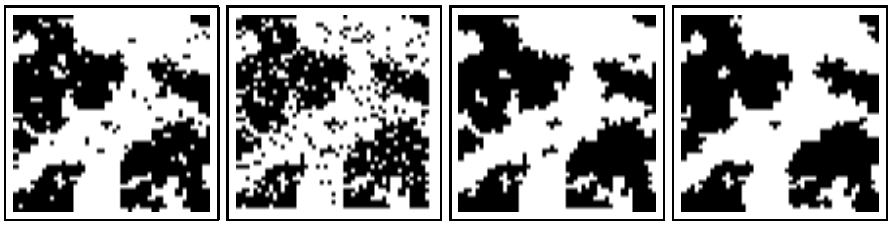

Fig. 1.19. From the left to the right, the original image ($T_s = 2.15$), the degraded image ($p_\tau = 0.1$), and typical restored images by the simulated annealing and the quantum annealing. The resultant bit-error rates are $p_b = 0.066400$ for the SA and $p_b = 0.058000$ for the QA.

1.6 Summary

In this article, we investigated the role of the quantum fluctuation introduced by means of the transverse field extensively. From the analysis of the infinite range model, we showed that the performances of the quantum MAP and MPM estimations are exactly the same as those of the thermal one. We derived the Nishimori-Wong condition on the effective amplitude of the transverse field and this information might be useful to determine the optimal amplitude of the transverse field for a given degraded image data. We also investigated the tolerance of the Sourlas codes to the quantum uncertainties in the prior distribution and discussed the condition on which the error-less ferromagnetic phase exists. We found that the Shannon's bound is not violated by the quantum fluctuation in the limit of $p \to \infty$. The analytic results of the image restoration were checked by the quantum Markov chain Monte Carlo method. The results supported the analysis of the infinite range model finely.

I hope that the present work provides some useful information for deep understanding of the optimization method based on the quantum fluctuation which is essentially different mechanism from the thermal hill-climbing.

The author thanks Prof. Bikas K. Chakrabarti and Dr. Arnab Das for organizing the workshop *Quantum Annealing and Other Optimization Meth-*

ods. He also acknowledges Saha Institute of Nuclear Physics (SINP) for kind hospitality during his stay in Kolkata.

References

1. H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing : An Introduction*, (Oxford University Press 2001).
2. K. Tanaka, J. Phys. A : Math. Gen. **35**, R81 (2002).
3. G. Winkler, *Image analysis, Random fields, and Markov Chain Monte Carlo Methods : A Mathematical Introduction*, (Springer 2002).
4. J. M. Pryce and A. D. Bruce, J. Phys. A. Math. Gen. **28**, 511 (1995).
5. N. Sourlas, Nature **339**, 693(1989).
6. P. Ruján, Phys. Rev. Lett. **70**, 2968(1993).
7. H. Nishimori, J. Phys. Soc. of Japan **62**, 2973(1993).
8. H. Nishimori and K. Y. M. Wong, Phys. Rev. E **60**, 132 (1999).
9. D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792(1975).
10. Y. Kabashima and D. Saad, Europhys. Lett. **45**, 98 (1999).
11. S. Geman and D. Geman, IEEE Trans. Pattern Anal. and Math. Intell. **PAMI-6-6**, 721 (1984).
12. S. Kirkpatrick, C. D. Gelatt Jr. and M. P. Vecchi, Science **220**, 671 (1983).
13. K. Tanaka and T. Horiguchi, IEICE **J80-A-12**, 2217 (1997) (in Japanese).
14. P. Amara, D. Hsu and J. E. Straub, J. Phys. Chem. **97**, 6715 (1993).
15. A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson and J. D. Doll, Chem. Phys. Lett. **219**, 343 (1994).
16. T. Kadowaki and H. Nishimori, Phys. Rev. E **58**, 5355 (1998).
17. E. Santoro, R. Martonak, E. Tosatti and R. Car, Science **295**, 2427 (2002).
18. C. Zener, Proc. R. Soc. London, Ser. A **137**, 696 (1932).
19. S. Miyashita, J. Phys. Soc. Jpn. **64**, 3207 (1995).
20. S. Miyashita, J. Phys. Soc. Jpn. **65**, 2734 (1996).
21. J. Inoue, Phys. Rev. E **63**, 046114 (2001).
22. B. K. Chakrabarti, A. Dutta and P. Sen, *Quantum Ising Phases ans Transitons in Transverse Ising Models*, (Springer 1995).
23. Y. Y. Goldshmidt, Phys. Rev. B **41**, 4858 (1990).
24. M. Suzuki, Prog. Theor. Phys. **58**, 1151 (1977).
25. H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
26. R. B. Ash, *Information Theory*, (Dover New York 1965).
27. J. Inoue and K. Tanaka, Phys. Rev. E **65**, 016125 (2002).
28. J. Inoue, APS March Meeting in Los Angels 2005, in preparation.