Interdisciplinary Information Sciences, Vol. 13, No. 1, pp. 43–48 (2007)

LDPC Decoding Dynamics from a PCA Viewpoint

Shinpei HARA[1*], Yuta AKIRA[1†], Eisuke ISHII[1‡], Masato INOUE[1§] and Masato OKADA[2,3¶]

[1]*Department of Electrical Engineering and Bioscience, School of Science and Engineering, Waseda University, Tokyo 169-8555, Japan*
[2]*Department of Complexity Science and Engineering, Graduate School of Frontier Sciences, The University of Tokyo, Chiba 277-8562, Japan*
[3]*RIKEN Brain Science Institute, Saitama 351-0198, Japan*

Received October 31, 2006; final version accepted December 28, 2006

Low-density parity-check (LDPC) code has recently become of great interest. The statistical mechanics approach has been used to reveal some characteristics of LDPC in the thermodynamic limit. In this paper, we analyze this system for finite size rather than within the thermodynamic limit through a principal component analysis (PCA) approach. Specifically, both the decoding dynamics of belief propagation (BP) and the phases of the system are visualized and discussed. The result implies that the decoding dynamics roughly corresponds to the system temperature we introduced, and this system has several phases such as ferromagnetic, paramagnetic, and 1RSB spin-glass phases.

KEYWORDS: low-density parity-check code, principal component analysis, spin-glass phase, mixed phase, replica symmetry breaking

1. Introduction

Low-density parity-check (LDPC) code [1, 2] has recently become of interest not only in the field of communication engineering, but also in the fields of large-scale information processing and statistical mechanics. Generally, a high-dimensional system such as LDPC code presents a difficult problem, but the statistical mechanics approach has been successfully used to gain useful understanding within the thermodynamic limit [3–5]. On the other hand, the analysis of LDPC code as a finite system has lagged. We believe the principal component analysis (PCA) approach [6, 7] could be used to remedy this and is therefore worth applying.

In this paper, we focus on two aspects of LDPC code. One is the analysis of the dynamics of a decoding algorithm by a sum-product algorithm; in other words, by belief propagation (BP) [8]. The second is the phases of a multi-body interaction system that arise from the parity-check matrix and channel noise. Potentially, both aspects can be managed through a PCA approach. Through this research, we hope to gain a better understanding of the system and contribute to the development of better decoding algorithms.

2. Model

In this section, we define the system model we have investigated. We consider LDPC code transmitted over a binary symmetric channel (BSC). We also consider the uniform prior of codewords for the prior distribution of a transmitted signal. We can then easily obtain the posterior distribution of the transmitted signal by Bayesian inference. Here, for the prior distribution, we introduce a small trick to make a PCA view more valuable; that is, the parity checks are not rigorous, but are relaxed according to temperature.

The prior distribution of transmitted signal $\boldsymbol{x} \in \{+1, -1\}^N$ is formally defined as

$$P(\boldsymbol{x}) \propto \prod_{\mu}^{M} \left(1 + \prod_{l \in \mathcal{M}(\mu)} x_l \right), \tag{1}$$

where $\mu = 1, \ldots, M$ denotes the parity index and $\mathcal{M}(\mu)$ denotes the set of bit indices involved in the μ-th parity. Similarly, $l = 1, \ldots, N$ denotes the bit index and $\mathcal{L}(l)$ denotes the set of parity indices linking to the l-th bit. $|\mathcal{M}(\mu)|$ and $|\mathcal{L}(l)|$ denote the degree of μ-th parity and the l-th bit, respectively. The proportion means the normalization of a probability function—*i.e.*, the summation of the probability for all possible arguments $\boldsymbol{x}$—should be 1. Throughout this paper, we adopt $\{+1, -1\}$ rather than $\{0, 1\}$ for binary expression. Every vector denotes a column vector unless otherwise stated.

* E-mail: shinpei-w@toki.waseda.jp
† E-mail: mazzantini@ruri.waseda.jp
‡ E-mail: cactus@ruri.waseda.jp
§ E-mail: masato.inoue@eb.waseda.ac.jp, corresponding author
¶ E-mail: okada@k.u-tokyo.ac.jp

Here, we introduce a small trick, a relaxation term, into the prior distribution,

$$P(\boldsymbol{x}) \propto \prod_{\mu}^{M}\left(1 + (\tanh\beta)\prod_{l\in\mathcal{M}(\mu)} x_l\right), \tag{2}$$

where $\beta \geq 0$. If $\beta = +\infty$, this prior distribution coincides exactly with the original one, so all the parity checks should be rigorously satisfied in transmitted signal $\boldsymbol{x}$. If $\beta = 0$, the restriction that the transmitted signal must be one of the codewords is completely ignored. We introduce this trick with $0 \leq \beta < +\infty$ to ensure reachability to the whole space through repetitive single bit flips in the conventional sequential Monte Carlo method.

The binary symmetric channel we use is defined as

$$P(\boldsymbol{y}|\boldsymbol{x}) = \prod_{l}^{N} \rho^{\frac{1-x_l y_l}{2}}(1-\rho)^{\frac{1+x_l y_l}{2}}, \tag{3}$$

where ρ is the bit flip rate, and $\boldsymbol{y}$ denotes the received signal. The posterior distribution is then given as

$$P(\boldsymbol{x}|\boldsymbol{y}) \propto \exp\left(\beta\sum_{\mu}^{M}\prod_{l\in\mathcal{M}(\mu)} x_l + \xi\sum_{l}^{N} y_l x_l\right). \tag{4}$$

where $\xi \equiv \frac{1}{2}\ln\frac{1-\rho}{\rho}$. Though not shown in this paper, the model can be easily extended to other memory-less channels, such as an additive white Gaussian channel. From here on, we call β the reciprocal temperature. We sometimes use temperature: $T \equiv 1/\beta$. We try to analyze both the decoding dynamics and phases of the system using this distribution.

3. Method

In this section, we explain the method to construct a PCA view for both the decoding dynamics and phases of the system. As the basic ideas of the PCA view have been reported in detail [7], we focus on the additional portion of the method. Briefly, we first constructed a parity-check matrix and transmitted one of the codewords through the channel. The received noisy signal $\boldsymbol{y}$ was used as one of the parameters of the posterior distribution. The empirical distribution of the posterior distribution was obtained by the exchange Monte Carlo (MC) method [9]. We then extracted the principal component (PC) vectors by applying PCA to the covariance matrix of the empirical distribution. The empirical distribution was re-used for projection onto the plane spanned by three types of axis: the PC vectors, $\boldsymbol{y}$, or the number of unsatisfied parity checks. As the decoding method here, we adopted BP. The time-course of the BP decoding dynamics was collected, and was also projected onto the same plane. Note that this method can be used without knowing the correct transmitted codewords (we used the *answer* only for drawing the *correct* codeword in purple double circle in Figs. 1–4).

3.1 Exchange Monte Carlo simulation

We sampled the spin configuration S times through equilibrium MC simulation, and obtained an empirical distribution to approximate the exact posterior distribution:

$$P(\boldsymbol{x}) \simeq \frac{1}{S}\sum_{s}^{S} I(\boldsymbol{x}^{(s)} = \boldsymbol{x}), \tag{5}$$

where $I()$ denotes the indicator function which returns 1 if the given condition is true, or 0 otherwise, and superscript (s) denotes the index of each sampled bit configuration.

To sample the bit configuration efficiently, we used the exchange MC method, which is also called parallel tempering. The exchange MC method prepares L systems that are identical except for the temperature of each system. One MC step (MCS) consists of N trials of the single spin flip and a succeeding set of temperature exchange trials at every neighboring temperature. These trials necessarily preserve the detailed balance of the L systems. Specifically, we used the Metropolis method; *i.e.*, the probability of the bit flip was defined as

$$\min\left(\frac{P(\boldsymbol{x}_i';\beta_i)}{P(\boldsymbol{x}_i;\beta_i)}, 1\right), \tag{6}$$

where $\boldsymbol{x}_i$, $\boldsymbol{x}_i'$, and β_i denote the bit configuration, the one with a single bit flip, and the reciprocal temperature of the ith corresponding system, respectively. The probability of the temperature exchange is defined as

$$\min\left(\frac{P(\boldsymbol{x}_j;\beta_i)P(\boldsymbol{x}_i;\beta_j)}{P(\boldsymbol{x}_i;\beta_i)P(\boldsymbol{x}_j;\beta_j)}, 1\right). \tag{7}$$

This temperature-exchange process is expected to accelerate the system relaxation, and the set of equilibrium configurations is efficiently obtained at each of the L temperatures in a single run.

3.2 Principal component analysis (PCA)

PCA can be done as follows. PCA diagonalizes the variance-covariance matrix of a given multivariate probability, $\boldsymbol{\Sigma} \equiv \mathrm{Var}[\boldsymbol{x}]$, as

$$\boldsymbol{\Sigma} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^T, \tag{8}$$

where ^{T} denotes the transpose, $\boldsymbol{V} \equiv [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N]$ is an $N \times N$ orthogonal matrix (*i.e.*, $\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}$, where $\boldsymbol{I}$ denotes the unit matrix), and $\boldsymbol{D} \equiv \mathrm{diag}[d_1, \ldots, d_N]$ ($d_1 \geq d_2 \geq \ldots \geq d_N \geq 0$) is an $N \times N$ diagonal matrix. $\boldsymbol{v}_i$ is the ith eigenvector (called the ith PC vector). The exact variance-covariance matrix of $\boldsymbol{x}$ is usually difficult to calculate, but we can estimate it from the empirical spin distribution of Eq. (5) as

$$\boldsymbol{\Sigma} \simeq \frac{1}{S}\sum_s^S (\boldsymbol{x}^{(s)} - \bar{\boldsymbol{x}})(\boldsymbol{x}^{(s)} - \bar{\boldsymbol{x}})^T, \tag{9}$$

where $\bar{\boldsymbol{x}} \equiv \frac{1}{S}\sum_s^S \boldsymbol{x}^{(s)}$ is the sample mean.

Here, we added a small modification to the original PCA method above. Specifically, we replaced the sampled spin configurations $\boldsymbol{x}^{(t)}$ with $R\boldsymbol{x}^{(t)}$, where

$$\boldsymbol{R} \equiv \boldsymbol{I} - \left(\frac{\boldsymbol{y}}{\sqrt{N}}\right)\left(\frac{\boldsymbol{y}}{\sqrt{N}}\right)^T, \tag{10}$$

is an $N \times N$ orthogonal projection matrix which removes the component along the $\boldsymbol{y}$ direction. This is done because we consider the $\boldsymbol{y}$ direction to have a special effect as the bias or external field in Eq. (4). This modification means the covariance matrix $\boldsymbol{\Sigma}$ is replaced with $\boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{R}$, and PCA tries to extract the PC vectors in the $N-1$ subspace, generally resulting in $\boldsymbol{v}_N = \pm\frac{\boldsymbol{y}}{\sqrt{N}}$, $d_N = 0$.

3.3 Visualization

We visualized the N-dimensional distribution using PC vectors, $\boldsymbol{y}$, or parity-check errors. Specifically, for PC vectors we chose the two PC vectors having the largest eigenvalues. Intuitively, the distribution spreads most widely along these vectors, and we can observe the shape of the high-dimensional distribution from the projection onto the plane spanned by these two vectors.

The visualization method utilizes the PC of each sampled spin configuration. Specifically, the mapping point of the sth sampled spin configuration is defined as

$$V_i^{(s)} \equiv \frac{\boldsymbol{x}^{(s)}}{\sqrt{N}} \cdot \boldsymbol{v}_i, \tag{11}$$

$$Y^{(s)} \equiv \frac{\boldsymbol{x}^{(s)}}{\sqrt{N}} \cdot \frac{\boldsymbol{y}}{\sqrt{N}}, \tag{12}$$

$$P^{(s)} \equiv \frac{1}{N}\sum_\mu^M \prod_{l\in\mathcal{M}(\mu)} x_l^{(s)}. \tag{13}$$

V_i and Y are normalized inner products. In accordance with the quantized coordinates, we constructed a frequency distribution map of the empirical spin distribution.

3.4 Belief propagation (BP)

For the decoding algorithm, we choose BP, which tries to infer the maximum posterior marginal (MPM) solution,

$$\hat{x}_l = \arg\max_{x_l} \sum_{\boldsymbol{x}_{\backslash l}} P(\boldsymbol{x}|\boldsymbol{y}), \tag{14}$$

using approximation, where $\backslash$ denotes the exclusion of the specified index. Specifically, BP iterates the following substitution and stops if the inferred $\hat{\boldsymbol{x}}$ reaches one of the codewords:

$$\eta_{\mu l}^{(t+1)} \equiv \xi y_l + \sum_{\mu'\in\mathcal{L}(l)\backslash\mu} \mathrm{atanh} \prod_{l'\in\mathcal{M}(\mu')\backslash l} \tanh\eta_{\mu' l'}^{(t)}, \tag{15}$$

$$h_l^{(t+1)} \equiv \xi y_l + \sum_{\mu'\in\mathcal{L}(l)} \mathrm{atanh} \prod_{l'\in\mathcal{M}(\mu')\backslash l} \tanh\eta_{\mu' l'}^{(t)}, \tag{16}$$

$$\hat{x}_l^{(t+1)} \equiv \mathrm{sign}\, h_l^{(t+1)}, \tag{17}$$

where we set $\eta_{\mu l}^{(0)} = h_l^{(0)} = \xi y_l$.

With respect to the projection coordinates of Eqs. (11)–(13), we replaced $x_l^{(s)}$ with $\tanh h_l^{(t)}$ for V_i and Y, and replaced $x_l^{(s)}$ with $\hat{x}_l^{(t)}$ for P.

4. Results and Discussion

In this section, we show results obtained through the above method and discuss these results. The preliminary results in the first part somewhat justify the model and the method. In the second part, we focus on the dynamics of BP decoding. In the third part, we focus on the phases of the system.

The parity-check matrix we used was made according to Gallager's construction. Specifically, we used 288-bit (3,6) regular LDPC code, in which $|\mathcal{L}(l)| = 3$ and $|\mathcal{M}(\mu)| = 6$. The error rate of BSC varied from $\rho = 0.08$ to 0.11. As the rate of this code is 0.5 (to be exact, it is equal to or greater than 146/288), it coincides with the channel capacity when $\rho = 0.110$. In our trials, BP decoded correctly when $\rho = 0.08$ and did not when $\rho = 0.09$, 0.10, or 0.11.

4.1 Bias and parity-check terms

The results shown in this subsection do not depend on the model, MC simulation, or PCA, and gives us one reason for constructing the model of Eq. (4) and using the projection matrix $\boldsymbol{R}$ of Eq. (10).

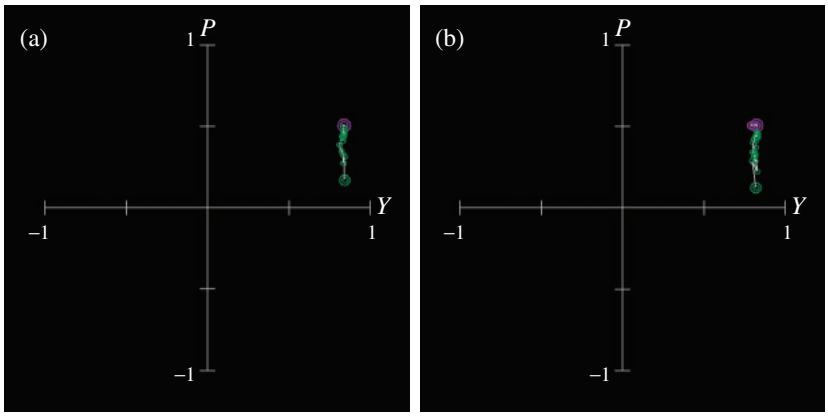

Fig. 1. Decoding dynamics of BP for the successful decoding case ($\rho = 0.08$) (a), and unsuccessful decoding case ($\rho = 0.09$) (b). Green circles denote decoding steps, and the double circle denotes the initial step. The purple double circle denotes the correct codeword. In both cases, Y varied very little.

Figure 1 shows the BP dynamics measured by the axes of the received vector (Y) and the number of unsatisfied parity checks (P). Empirically, the Y value does not vary very much regardless of whether decoding is successful or unsuccessful. This tells us one qualitative story of BP dynamics: that the second term (the bias term) in the exponent of Eq. (4) remains constant during decoding, and the first term (the parity-check term) gradually increases. This implies one hypothesis that the model expresses both the early steps of BP with low β and converging steps with high β. Furthermore, the direction of $\boldsymbol{y}$ can be considered less important in the PCA approach, so we explicitly excluded the component along the direction $\boldsymbol{y}$ using the projection matrix $\boldsymbol{R}$.

4.2 Decoding dynamics

Figures 2(a) and (b) shows successful decoding dynamics in the $V_1 - P$ and $V_1 - V_2$ planes. As expected from the above, the early steps of BP dynamics are consistent with the distribution with high temperature and the converging steps match the distribution with low temperature. The $V_1 - V_2$ and $V_1 - V_2$ plane views also explain the dynamics as shown in Figs. 2(c) and (d).

Figure 3 shows unsuccessful decoding cases, where the plane view can be considered inadequate for analyzing dynamics. This is because, in such cases, more than one codeword are usually strong candidates and they are linearly independent of each other. If the number of such candidates exceeds three, the two-dimensional plane view is not enough to estimate the distances between the candidates or between them and the inferred vector. Nonetheless, Fig. 3 gives us some information; the inferred vector seems to have initially moved towards one of the incorrect codewords and then turned around. After that, it appears to have passed through another incorrect codeword, but, to be exact, it did not reach any of the codewords; *i.e.*, the 'depth' of the inferred vector actually differed from that of the codeword.

4.3 Phases

The model can be understood as a multi-body interaction system whose bonds are defined by a parity-check matrix and whose external fields are defined by noise [10]. In this system, the number of the restrictions, M, is less than the number of spins, N. This system is therefore sparse and a non-frustrating system if external fields do not exist or are weak. In such systems, the ferromagnetic phase is usually dominant, which can be roughly considered a successful

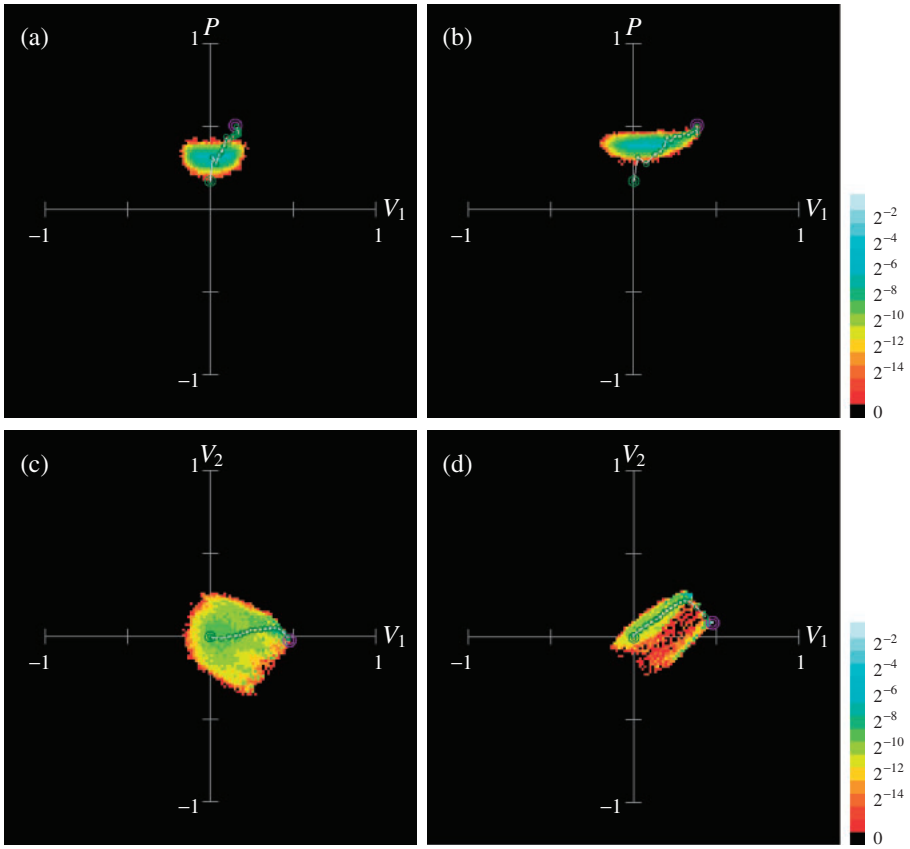

Fig. 2. Distribution maps with successful decoding dynamics ($\rho = 0.08$): (a) $T = 1.5$, (b) $T = 1.0$, (c) $T = 0.8$, and (d) $T = 0.7$.

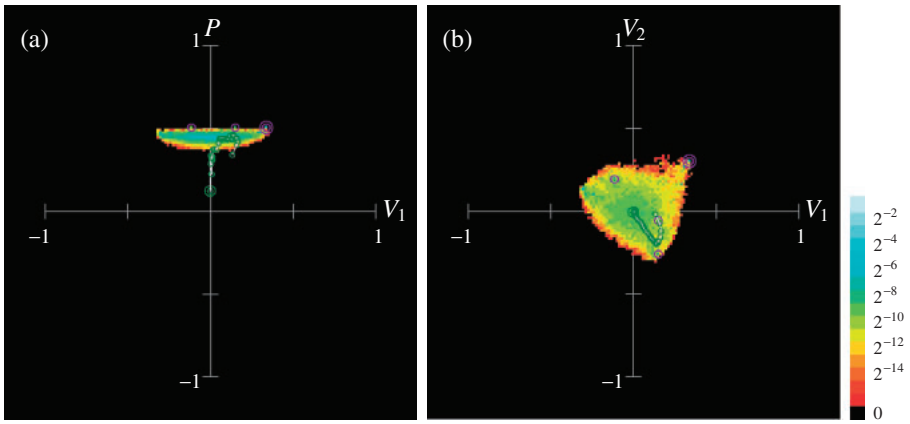

Fig. 3. Distribution maps with unsuccessful decoding dynamics ($\rho = 0.09$): (a) $T = 0.7$ and (b) $T = 0.7$.

decoding situation. On the other hand, the system is frustrated with strong external fields because the signs of bond strength are mixed after appropriate gauge transformation. In such systems, a spin-glass phase can be considered dominant and can correspond to an unsuccessful decoding situation. In both systems, the paramagnetic phase becomes dominant when the temperature is high, and this might be considered the early steps of decoding.

It is possible that these phases can be distinguished using a PCA approach [7]. Indeed, these phases are distinguishable as shown in Fig. 4. With respect to the ferromagnetic phase, we could not draw the PCA view because the bits remained fixed to the correct codeword rather than forming a distribution. This feature implies the ferromagnetic phase. With respect to the mixed phase, Figs. 4(a) and (b) show characteristic triangle formation, implying a one-step replica symmetry breaking (1RSB) spin-glass phase [6, 7], but the triangle is smaller than that in the reference. Furthermore, the triangle size seems to depend on the strength of the external field, or equivalently, channel noise. Since the average direction of the spin configuration is not the zero vector (the component along which the bias direction exists) in this system, it might be a 1RSB mixed phase; however, this needs further investigation. With respect to the paramagnetic phase, Figs. 4(c) and (d) each show a concentric circle around the origin, which

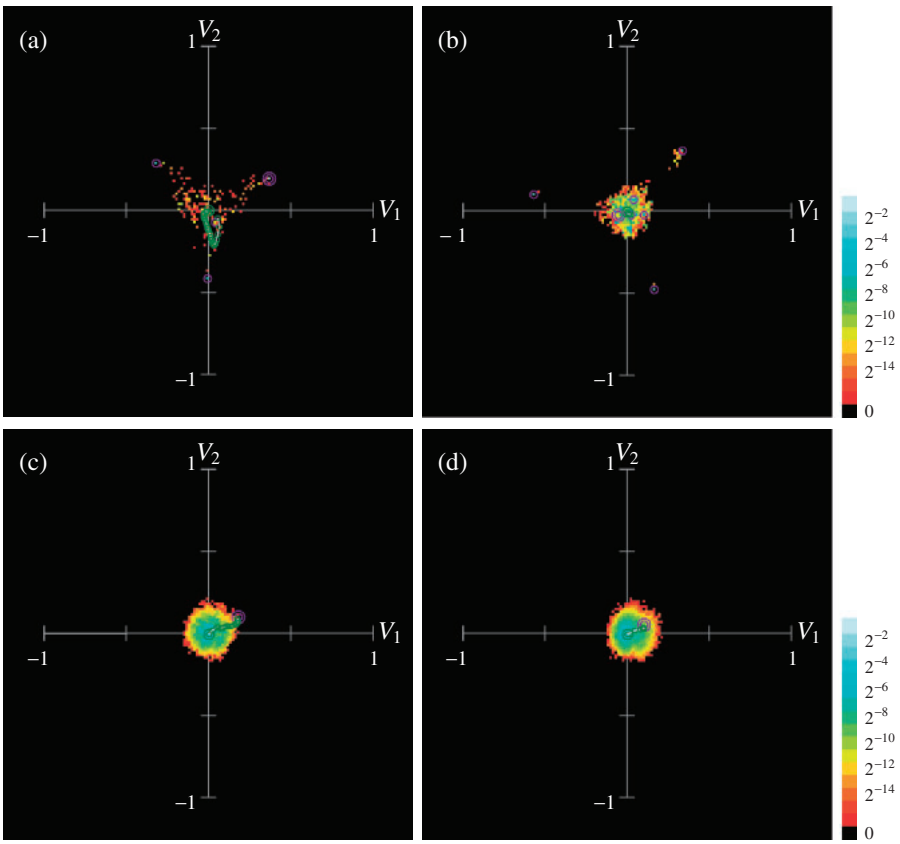

Fig. 4. Distribution maps: (a) unsuccessful case ($\rho = 0.09$) with $T = 0.4$, (b) unsuccessful case ($\rho = 0.11$) with $T = 0.4$, (c) successful case ($\rho = 0.08$) with $T = 4.0$, and (d) unsuccessful case ($\rho = 0.11$) with $T = 4.0$.

implies the paramagnetic phase. Also in this case, the average direction of the spin configuration is not the zero vector, so it might not be appropriate to call this the paramagnetic phase.

5. Conclusion

We analyzed the system of an LDPC decoding problem through a principal component analysis (PCA) view. The decoding dynamics of belief propagation (BP) roughly corresponded to the distribution of the model in both early steps with high temperature and converging steps with low temperature. In addition, the PCA result implies that the system has ferromagnetic, paramagnetic, and 1RSB spin-glass phases if we ignore the bias or the external field direction, but this needs further investigation.

Acknowledgements

This work was partially supported by Grant-in-Aid for Scientific Research on Priority Areas No. 18079012, Grant-in-Aid for Scientific Research (B) No. 17340116, and Waseda University Grant for Special Research Projects.

REFERENCES

[1] Gallager, R. G., *Low density parity check codes*, in Research Monograph series, Cambridge, MIT Press, 1963.
[2] MacKay, D. J. C., *IEEE Trans. Inform. Theory*, **IT-45**(2): 399 (1999).
[3] Kabashima, Y., and Saad, D., *Europhys. Lett.*, **44**: 668 (1998).
[4] Kabashima, Y., *J. Phys. Soc. Jpn.*, **72**: 1645 (2003).
[5] Kabashima, Y., and Saad, D., *J. Phys. A*, **37**: R1–R43 (2004).
[6] Inoue, M., Hukushima, K., and Okada, M., *Prog. Theor. Phys. Suppl.*, **157**: 246 (2005).
[7] Inoue, M., Hukushima, K., and Okada, M., *J. Phys. Soc. Jpn.*, **75**: 084003 (2006).
[8] Pearl, J., *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, 1988.
[9] Hukushima, K., and Nemoto, K., *J. Phys. Soc. Jpn.*, **65**: 1604 (1996).
[10] Nishimori, H., *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press Inc., New York, 2001.