交換モンテカルロ法における熱浴型交換率の解析

永田　賢二[†]　　渡辺　澄夫[††]

† 東京工業大学 総合理工学研究科 知能システム科学専攻 〒 226-8503 横浜市緑区長津田 4259
†† 東京工業大学 精密工学研究所 〒 226-8503 横浜市緑区長津田 4259
E-mail: †kenji.nagata@cs.pi.titech.ac.jp, ††swatanab@pi.titech.ac.jp

あらまし　交換モンテカルロ法は、MCMC 法の改良アルゴリズムとして様々な分野でその有効性が示されている。しかしながら、交換モンテカルロ法の数学的な性質は未だ明らかにされていないために、理論的に最適な交換モンテカルロ法の設計法は確立されていない。我々は、先行研究において、任意の確率分布の低温極限におけるメトロポリス型における平均交換率と対称カルバック距離の漸近形を解析的に導出することで両者の関係を明らかにし、最適な温度パラメータの設定法を提案した。本研究では、低温極限における熱浴型における平均交換率の漸近挙動を明らかにし、メトロポリス型との比較により、両者の数学的な性質の相違を明らかにする。
キーワード　マルコフ連鎖モンテカルロ法、交換モンテカルロ法、カルバック距離、交換率、平均交換率

Analysis of Exchange Ratio for Heat Bath Type in Exchange Monte Carlo Method

Kenji NAGATA[†] and Sumio WATANABE[††]

† Department of Computer Science Tokyo Institute of Technology, 4259
Nagatsuda,Midori-ku,Yokohama,226-8503 Japan
†† PI Lab., Tokyo Institute of Technology.
E-mail: †kenji.nagata@cs.pi.titech.ac.jp, ††swatanab@pi.titech.ac.jp

Abstract　The exchange Monte Carlo method is well known as an improved algorithm of Markov Chain Monte Carlo method. Although its effectiveness has been shown in many fields, the mathmatical foundation of exchange Monte Carlo method has not yet been established. In our previous work, we analytically clarify the asymptotic behavior of symmetirized Kullback divergence and exchange ratio for Metropolis type in low temperature limit. In this paper, we analytically charify the asymptotic behavior of the exchange ratio for heat bath type in low temperature limit, and discuss the mathematical property of exchange ratio for Metropolis type and for heat bath type.
Key words　Markov Chain Monte Carlo Method, Exchange Monte Carlo Method, Kullback Divergence, Exchange Ratio, Average Exchange Ratio

1. Introduction

The Markov Chain Monte Carlo (MCMC) method is well known as an algorithm to generate the sample sequence which converges to a target distribution. This algorithm is widely used in many fields such as statistics, bioinformatics and information science. However, it requires huge computational cost to generate the sample sequence, in particular, in the case that the target distribution has high potential barriers [6] and that the ground state of target distribution is not one point but an analytic set [11].

Recently, various improvements of MCMC method have been developed based on the idea of extended ensemble methods, which are surveyed in [7]. These methods give us a general strategy to overcome the problem of huge computational cost. An exchange Monte Carlo (MC) method is well known as one of the extended ensemble methods [6]. This method is to generate the sample sequence from a joint distribution, which consists of many distributions with different temperatures. Its algorithm is based on two steps of MCMC simulations. One is the conventional update of MCMC simulation for each distribution. The other is the exchange process between two sequences with a certain probability. The effectiveness of this algorithm has been shown in an optimization problem [5] [12], a protein-folding problem [13] and the Bayesian learning in hierarchical learning machines [10].

When we design the exchange MC method, the setting of temperature is very important [4]. The values of temperature have close relation to the exchange ratio and its average, with which the exchange between two sequence is accepted. In order to make the exchange MC method efficient, the exchange ratio needs to be not low and not too high. Therefore, the optimal setting of temperature enables us to design the efficient exchange MC method. The symmetrized Kullback divergence between two distributions with different temperature is used as a criterion for the setting of temperature because this Kullback divergence has relation to the average exchange ratio [7]. Based on this fact, the design method for setting of temperature has been proposed. However, this method needs some previous simulations. Moreover, the mathematical relation between the symmetrized Kullback divergence and the average exchange ratio has not been clarified.

In our previous work, we clarified the asymptotic behavior of the symmetrized Kullback divergence and the average exchange ratio for Metropolis type, and discussed the property of two functions. Moreover, based on this result, the optimal setting for exchange MC method was proposed [8] [9].

In this paper, we clarify the asymptotic behavior of the average exchange ratio for heat bath type, and discuss the property of exchange ratio for Metropolis type and for heat bath type.

This paper consists of six chapters. In Chapter 2, we explain the framework of exchange MC method and the design of exchange MC method. Our previous work for exchange MC method is described in Chapter 3. In Chapter 4, the main result of this paper is described. Discussion and Conclusion are followed in Chapter 5 and 6.

2. Background

2.1 Exchange Monte Carlo Method

Suppose that $w \in R^d$ and our aim is to generate the sample sequence from the following target probability distribution with a energy function $H(w)$ and a probability distribution $\varphi(w)$,

$$p(w) = \frac{1}{Z(n)} \exp(-nH(w))\varphi(w),$$

where $Z(n)$ is the normalization constant. The exchange MC method treats a compound system which consists of non-interacting K sample sequences of the system concerned. The k-th sample sequence $\{w_k\}$ converges in law to the random variable which is subject to the following probability distribution,

$$p(w|t_k) = \frac{1}{Z(nt_k)} \exp(-nt_k H(w))\varphi(w) \quad (1 \leqq l \leqq K),$$

where $t_1 < t_2 < \cdots < t_K$. Given a set of the temperatures $\{t\} = \{t_1, \cdots, t_K\}$, the joint distribution for finding $\{w\} = \{w_1, w_2, \cdots, w_K\}$ is expressed as a simple product formula,

$$p(\{w\}) = \prod_{k=1}^{K} p(w_k|t_k). \quad (1)$$

The exchange MC method is based on two types of updating in constructing a Markov chain. One is conventional updates of MCMC simulation such as Gibbs sampler and Metropolis algorithm for each target distribution $p(w_k|t_k)$. The other is the position exchange between two sequences, that is, $\{w_k, w_{k+1}\} \to \{w_{k+1}, w_k\}$. The transition probability u is determined by the detailed balance condition for the joint distribution of Eq.(1). We have two types of position exchange. One is the Metropolis type as follows,

$$\begin{aligned} u_1 &= \min(1, r), \\ r &= \frac{p(w_{k+1}|t_k)p(w_k|t_{k+1})}{p(w_k|t_k)p(w_{k+1}|t_{k+1})} \\ &= \exp(n(t_{k+1} - t_k)(H(w_{k+1}) - H(w_k))). \end{aligned} \quad (2)$$

The other is the heat bath type,

$$\begin{aligned} u_2 &= \frac{p(w_{k+1}|t_k)p(w_k|t_{k+1})}{p(w_k|t_k)p(w_{k+1}|t_{k+1}) + p(w_{k+1}|t_k)p(w_k|t_{k+1}} \\ &= \frac{1}{2}\left\{1 + \tanh\left(\frac{n}{2}(t_{k+1} - t_k)\left(H(w_{k+1}) - H(w_k)\right)\right)\right\} \end{aligned} (3)$$

Hereafter, we call u_1 and u_2 exchange ratios. Under these updates, the joint distribution of Eq.(1) is invariant because these updates satisfy the detailed balance condition for the distribution of Eq.(1) [6].

Consequently, the following two steps are carried out alternately:

(1) Each sequence is generated simultaneously and independently for a few iteration by conventional MCMC method.

(2) Two positions are exchanged with the exchange ratio u_1 or u_2.

The advantage of exchange MC method is to make the convergence of sample sequence earlier comparing the conventional MCMC method. The conventional MCMC method requires huge computational cost to generate sample sequence from the target distribution. The reason is that the sample sequence is hard to converge in law to the random variable which is subject to the target distribution, in particular, in the case that the target distribution has high potential barriers, and that the ground state of target distribution is not one point but an analytic set. The exchange MC method can realize the efficient sampling by preparing a simple distribution such as a normal distribution, which is easy for sample sequence to converge. In practical, we set the temperature of target distribution as $t_K = 1$, and that of simple distribution as $t_1 = 0$.

2. 2 Design of Exchange Monte Carlo Method

When we design the exchange MC method, the setting of temperature is important to make the exchange MC method efficient. As we can see in Eq.(2) and Eq.(3), temperature has close relation to the exchange ratio. Therefore, temperature is a important parameter in adjusting the exchange ratio and its average.

For the efficient exchange MC method, the set of temperature needs to optimize so that the average exchange ratio for two adjacent distribution become not low and not too high. In order to carry out the efficient exchange MC method, the time for a sample to move from end to end (from t_1 to t_K) in the space of temperature is good to be short. Therefore, it is not efficient for the average exchange ratio to be low. On the contrary, to make the average exchange ratio high, the range of temperature has to be very small, that is, the total number K of temperature has to be large. Therefore, this setting is not also efficient because it needs huge cost to generate the sample from each distribution. In practical, the set of temperature is configured so that the average exchange ratio becomes equal for all combinations of distributions.

As the criterion for setting of temperature, The following symmetrized Kullback divergence $I(t_k, t_{k+1})$ is used [7],

$$I(t_k, t_{k+1}) = \int p(w_k|t_k) \log \frac{p(w_k|t_k)}{p(w_k|t_{k+1})} dw_k + \int p(w_{k+1}|t_{k+1}) \log \frac{p(w_{k+1}|t_{k+1})}{p(w_{k+1}|t_k)} dw_{k+1}.$$

This function has the following property,

$$E[\log r] = -I(t_k, t_{k+1}),$$

where $E[\log r]$ means the average of $\log r$ over the joint distribution $p(w_k|t_k) \times p(w_{k+1}|t_{k+1})$. Moreover, when the free energy $F(nt)$ is defined as follows,

$$F(nt) = -\log \int \exp(-ntH(w))\varphi(w)dw,$$

the following equation is satisfied in small range of temperature,

$$I(t_k, t_{k+1}) = \frac{\partial^2 F}{\partial t^2}\Big|_{t=t_k} (t_{k+1} - t_k)^2.$$

Hence, the set of temperature can be set so that the symmetrized Kullback divergence becomes constant by setting the range of temperature in inverse proportion to $\sqrt{\partial^2 F/\partial t^2}$. However, the mathematic definition of average exchange ratio $J(t_k, t_{k+1})$ is as follows,

$$J_1(t_k, t_{k+1}) = \int\int u_1 P(w_k|t_k)P(w_{k+1}|t_{k+1})dw_k dw_{k+1}$$

$$J_2(t_k, t_{k+1}) = \int\int u_2 P(w_k|t_k)P(w_{k+1}|t_{k+1})dw_k dw_{k+1},$$

whose properties are not clarified. Therefore, the mathematical relation between the symmetrized Kullback divergence and the average exchange ratio had not been analytically clarified.

3. Our Previous Work

In our previous work, we clarify the asymptotic behavior of symmetrized Kullback divergence and average exchange ratio for Metropolis type in low temperature limit [8] [9].

we consider the exchange MC method between the following two distributions,

$$p_1(w) = \frac{1}{Z(nt)} \exp(-ntH(w))\varphi(w)$$

$$p_2(w) = \frac{1}{Z(n(t+\Delta t))} \exp(-n(t+\Delta t)H(w))\varphi(w).$$

For these distributions, the symmetrized Kullback divergence I and the average exchange ratio J_1 for Metropolis type are rewritten as follows,

$$I = \int p_1(w_1) \log \frac{p_1(w_1)}{p_2(w_1)} dw_1 + \int p_2(w_2) \log \frac{p_2(w_2)}{p_1(w_2)} dw_2$$

$$J_1 = \int\int u_1 p_1(w_1)p_2(w_2)dw_1 dw_2,$$

where u_1 is a function of w_1 and w_2 as Eq.(2).

In general, either the distribution $p_1(w)$ or $p_2(w)$ does not converge to the normal distribution as $n \to \infty$ because the hessian of energy function $H(w)$ is not positive definite. We can assume $H(w) \geqq 0$ and $H(w_0) = 0 (\exists w_0)$ without loss of generality. The zeta function of $H(w)$ and $\varphi(w)$ is defined by

$$\zeta(z) = \int H(w)^z \varphi(w)dw,$$

where z is a one complex variable. $\zeta(z)$ is a holomorphic function in the region of $Re(z) > 0$, and can be analytically continued to the meromorphic function on the entire complex plane, whose poles are all real, negative, and rational numbers [3] [15]. We also define the rational number $-\lambda$ as the largest pole of zeta function $\zeta(z)$ and the natural number m as its order.

Then, the following theorem was proven.

[Theorem 1] The symmetrized Kullback divergence I and the average exchange ratio J_1 for Metropolis type respectively converge to the following equations as $n \to \infty$,

$$I \to \lambda \left(\frac{\Delta t}{t}\right)^2 \left(1 - \frac{\Delta t}{t} + O\left(\left(\frac{\Delta t}{t}\right)^2\right)\right)$$

$$J_1 \to 1 - \frac{|\Delta t|}{t} \frac{\Gamma(\lambda + \frac{1}{2})}{\sqrt{\pi}\Gamma(\lambda)} + O\left(\left(\frac{\Delta t}{t}\right)^2\right).$$

From this theorem, the symmetrized Kullback divergence I and the average exchange ratio J_1 for Metropolis type become constant by the temperature setting that the value $\frac{\Delta t}{t}$ becomes constant, that is, the set $\{t_k\}$ of temperature is set as geometrical progression. Moreover, we can see the difference that the average exchange ratio is expressed by the liear function of Δt which has the maximum value 1 at $\Delta t = 0$, while the symmetrized Kullback divergence is the quadratic which has the minimum value 0 at $\Delta t = 0$.

In this paper, we analytically clarify the asymptotic behavior of average exchange ratio for heat bath type in low temperature limit, and compare the property of average exchange ratio for Metropolis type to that for heat bath type.

4. Main Result

In this section, we show the main result of this paper.

We analyze the following average exchange ratio J_2 for heat bath type,

$$J_2 = \int\int u_2 p_1(w_1) p_2(w_2) dw_1 dw_2,$$

where u_2 is function of w_1 and w_2 as we can see in Eq.(3).

[Theorem 2] The average exchange ratio J_2 for heat bath type converges to the following equation as $n \to \infty$,

$$J_2 \to \frac{1}{2}\left(1 - \frac{\lambda}{2}\left(\frac{\Delta t}{t}\right)^2 + O\left(\left(\frac{\Delta t}{t}\right)^2\right)\right).$$

(proof of Theorem 2)The average exchange ratio J_2 is expressed by the definition of exchange ratio u_2 as follows,

$$J_2 = \frac{1}{2}\left\{1 + \int dw_1 \int dw_2 \tanh\left(\frac{n\Delta t}{2}(H(w_2) - H(w_1))\right) \times \frac{e^{-ntH(w_1)}\varphi(w_1)}{Z(nt)} \frac{e^{-n(t+\Delta t)H(w_2)}\varphi(w_2)}{Z(n(t+\Delta t))}\right\}.$$

Hence, by defining

$$J_2^* = \int dw_1 \int dw_2 \tanh\left(\frac{n\Delta t}{2}(H(w_2) - H(w_1))\right) \times e^{-ntH(w_1)}\varphi(w_1) e^{-n(t+\Delta t)H(w_2)}\varphi(w_2),$$

we obtain

$$J_2 = \frac{1}{2}\left\{1 + \frac{J_2^*}{Z(nt)Z(n(t+\Delta t))}\right\}.$$

The function J_2^* is expressed by using the Dirac delta function as follows,

$$J_2^* = \int_0^\infty ds_1 \int_0^\infty ds_2 \tanh\left(\frac{n\Delta t}{2}(s_2 - s_1)\right) \times e^{-nts_1} e^{-n(t+\Delta t)s_2} \times \int dw_1 \delta(s_1 - H(w_1))\varphi(w_1) \times \int dw_2 \delta(s_2 - H(w_2))\varphi(w_2)$$

The integration for w in the above equation is well known as the state density function $V(s)$. This function has the following asymptotic form for $s \to 0$ [14] [15],

$$V(s) = \int \delta(s - H(w))\varphi(w) dw \cong cs^{\lambda-1}(-\log s)^{m-1},$$

where the real number $c > 0$ is the function of $H(w)$ and $\varphi(w)$. From this equation, the function J_2^* is given by,

$$J_2^* \cong \int_0^\infty ds_1 \int_0^\infty ds_2 \tanh\left(\frac{n\Delta t}{2}(s_2 - s_1)\right) e^{-nts_1} e^{-n(t+\Delta t)s_2} \times cs_1^{\lambda-1}(-\log s_1)^{m-1} cs_2^{\lambda-1}(-\log s_2)^{m-1}.$$

By putting $s_1' = nts_1$ and $s_2' = nts_2$, it follows that

$$\begin{aligned} J_2^* &\cong \int_0^\infty \frac{ds_1'}{nt} \int_0^\infty \frac{ds_2'}{nt} \tanh\left(\frac{1}{2}\frac{\Delta t}{t}(s_2' - s_1')\right) e^{-\frac{\Delta t}{t}s_2'} \\ &\quad \times e^{-s_1'}\left(\frac{s_1'}{nt}\right)^{\lambda-1}(\log nt - \log s_1')^{m-1} \\ &\quad \times e^{-s_2'}\left(\frac{s_2'}{nt}\right)^{\lambda-1}(\log nt - \log s_2')^{m-1} \\ &\cong \frac{c^2(\log nt)^{2(m-1)}}{(nt)^{2\lambda}}\left\{O\left(\left(\frac{\Delta t}{t}\right)^3\right)\right. \\ &\quad + \frac{1}{2}\frac{\Delta t}{t}\int_0^\infty ds_1' \int_0^\infty ds_2' e^{-s_1'} e^{-s_2'} \\ &\quad \times (s_1'^{\lambda-1} s_2'^{\lambda} - s_1'^{\lambda} s_2'^{\lambda-1}) \\ &\quad - \frac{1}{2}\left(\frac{\Delta t}{t}\right)^2 \int_0^\infty ds_1' \int_0^\infty ds_2' e^{-s_1'} e^{-s_2'} \\ &\quad \left. \times (s_1'^{\lambda-1} s_2'^{\lambda+1} - s_1'^{\lambda} s_2'^{\lambda})\right\} \\ &= \frac{c^2(\log nt)^{2(m-1)}}{(nt)^{2\lambda}} \\ &\quad \times \left(-\frac{\lambda\Gamma(\lambda)^2}{2}\left(\frac{\Delta t}{t}\right)^2 + O\left(\left(\frac{\Delta t}{t}\right)^3\right)\right). \end{aligned}$$

In the above analysis, we used the Taylor expansion of

$$f\left(\frac{\Delta t}{t}\right) = \tanh\left(\frac{1}{2}\frac{\Delta t}{t}(s_2' - s_1')\right) e^{-\frac{\Delta t}{t}s_2'}$$

around $\frac{\Delta t}{t} = 0$ as follows,

$$f\left(\frac{\Delta t}{t}\right) = \frac{1}{2}(s_2' - s_1')\frac{\Delta t}{t} - \frac{1}{2}s_2'(s_2' - s_1')\left(\frac{\Delta t}{t}\right)^2 + O\left(\left(\frac{\Delta t}{t}\right)^3\right).$$

Consequently, the average exchange ratio J_2 is equal to,

$$J_2 = \frac{1}{2}\left\{1 + \frac{J_2^*}{Z(nt)Z(n(t+\Delta t))}\right\} \to \frac{1}{2}\left\{1 - \frac{\lambda}{2}\left(\frac{\Delta t}{t}\right)^2 + O\left(\left(\frac{\Delta t}{t}\right)^3\right)\right\},$$

which completes this theorem.(Q.E.D)

Theorem 2 claims that the average exchange ratio for heat bath type is also constant by setting the value $\frac{\Delta t}{t}$ constant. Consequently, if we set the value $\frac{\Delta t}{t}$ constant, the symmetrized Kullback divergence and the average exchange ratio for both type are clarified to be constant. Moreover, we can see the property that the average exchange ratio for heat bath type is quadratic function of Δt which has the maximum value of $\frac{1}{2}$ at $\Delta t = 0$.

5. Discussion

In this paper, we analyzed the average exchange ratio for heat bath type in the low temperature limit, and clarified the mathematical relation among the symmetrized Kullback divergence and the average exchange ratios for two types. We have shown the following three properties,

（1） When the symmetrized Kullback divergence for arbitrary temperature t is constant, the average exchange ratio for each type is also constant.

（2） Then, the set of temperature $\{t_k\}$ is set as geometrical progression.

（3） The average exchange ratio for Metropolis type is expressed by the linear function, while the other two functions are the quadratic functions.

Let us discuss three points in association to this paper.

Firstly, we discuss the relation between the shape of target distribution and the setting of temperature. As we can see in Theorem 2, once the set of temperature $\{t_k\}$ is determined, the value of average exchange ratio depends on the value λ. The average exchange ratios for both type become small for the distribution with large value λ because the coefficient of linear term for the average exchange ratio is the monotonically increasing function for λ [8] [9]. Therefore, for the target distribution with small value λ, the exchange MC method works more efficiently. On the other hand, by comparing two distributions whose ground state are one point and analytic set in the sample space, the latter distribution is well known to have smaller value λ than the former distribution [2] [16]. Consequently, the exchange MC method can work efficiently for the target distribution with the energy function whose ground state is analytic set. As an example of such distributions, the Bayesian posterior distribution in hierarchical learning machines such as neural networks and normal mixtures is well known, and this theorem shows the availability of exchange MC method for the Bayesian learning in hierarchical learning machines.

Secondly, we discuss the properties of average exchange ratio for Metropolis type and for heat bath type. When we implement the EMC method, which type should be used as an exchange ratio, Metropolis type or heat bath type, often becomes a problem. By comparing each exchange ratio, the computational cost is almost equal. Therefore, we should select the type to make the average exchange ratio high for the same setting of temperature. According to Theorem 1 and 2, the average exchange ratio for Metropolis type is clarified to be higher than that for heat bath type in low temperature limit. This claims the effectiveness of Metropolis type for the EMC method.

Finally, we discuss the design of exchange MC method. This theorem gives us the design method of optimal temperature in order to make the average exchange ratio constant. However, the optimum value of average exchange ratio is not clarified, which leads to the design of optimal number K of temperature. Moreover, since the exchange MC method includes the algorithm of conventional MCMC method, the design of conventional MCMC method should be considered in the future.

6. Conclusion

In this paper, we analyzed the average exchange ratio for heat bath type in the low temperature limit, and clarified the mathematical relation among the symmetrized Kullback divergence and the average exchange ratios for two types. We have shown the following three properties,

（1） When the symmetrized Kullback divergence for arbitrary temperature t is constant, the average exchange ratio for each type is also constant.

（2） Then, the set of temperature $\{t_k\}$ is set as geometrical progression.

（3） The average exchange ratio for Metropolis type is expressed by the linear function, while the other two functions are the quadratic functions.

As the future works, verifying the theoretical result obtained by this study by some experiments, constructing the design of exchange MC method, and applying these results to the practical problems such as the Bayesian learning should be addressed.

7. Acknowledgement

This work was supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for JSPS Research Fellows 18-5809 and for Scientific Research 18009007.

文　　献

[1] H.Alzer, “On some inequalities for the Gamma and Psi functions,” Mathematics of computation, 66 (217), pp.373-389, 1997.

[2] M.Aoyagi, S.Watanabe, “Stochastic complexities of reduced rank regression in Bayesian estimation,” Neural Networks, Vol.18, No.7, pp.924-933, 2005.

[3] M. F. Atiyah, “Resolution of singularities and division of distributions,” Commun. Pure Appl. Math., vol. 13, pp. 145-150, 1970.

[4] K.Hukushima, “Domain-Wall Free-Energy of Spin Glass Models: Numerical Method and Boundary Conditions,” Physical Review E 60, pp.3606-3613, 1999.

[5] K.Hukushima, “Extended ensemble Monte Carlo approach to hardly relaxing problems,” Computer Physics Communications, 147, pp.77-82, 2002.

[6] K.Hukushima, K.Nemoto, “Exchange Monte Carlo Method and Application to Spin Glass Simulation,” Journal of the Physical Society of Japan, Vol.65, No.6, pp.1604-1608, 1996.

[7] Y.Iba, “Extended Ensemble Monte Carlo”, International Journal of Modern Physics, C12, pp.623-656, 2001.

[8] K.Nagata, S.Watanabe, “Analysis of Exchange Ratio for Exchange Monte Carlo Method,” Proc. of The First IEEE Symposium on Foundation of Computational Intelligence (FOCI’07), to appear.

[9] K.Nagata, S.Watanabe, “On Relation between Exchange Ratio and Kullback Divergence in Exchange Monte Carlo Method,” Technical Report of IEICE, Vol.106, No.279, pp.43-48, 2006.

[10] K.Nagata, S.Watanabe, “The Exchange Monte Carlo Method for Bayesian Learning in Singular Learning Machines,” Proc. of World Congress on Computational Intelligence 2006 (WCCI2006), pp.6383-6389, 2006.

[11] N.Nakano, K.Takahashi, S.Watanabe, “On the Evaluation Criterion of the MCMC Method in Singular Learning Machines” , Trans. of IEICE, Vol.J88-D-2, No.10, pp.2011-2020, 2005.

[12] K.Pinn, C.Wieczerkowski, “Number of magic squares from parallel tempering Monte Carlo,” Int. J. Mod. Phys. C9, 541, 1998.

[13] Y.Sugita, Y.Okamoto, “Replica-exchange molecular dynamics method for protain folding,” Chemical Physics Letters, 314, pp.141-151, 1999.

[14] S.Watanabe, “Algebraic analysis for nonidentifiable learning machines,” Neural Computation, Vol.13, No.4, pp.899-933, 2001.

[15] S.Watanabe, “Algebraic geometrical methods for hierarchical learning machines,” Neural Networks, Vol.14, No.8, pp.1049-1060, 2001.

[16] K.Yamazaki, S,Watanabe, “Singularities in mixture models and upper bounds of stochastic complexity,” Neural Networks, Vol.16, No.7, pp.1029-1038, 2003.