Adaptation and linear-response theory

Toyonori Munakata and Satoshi Oyama
Department of Applied Mathematics and Physics, Kyoto University, Kyoto 606, Japan
(Received 7 March 1996)

Adaptation, i.e., slow time variation of control parameters to achieve better performance, which is usually estimated by some object or cost function, is investigated based on linear-response theory. Adaptive Monte Carlo methods such as entropy (or multicanonical) sampling and nonsupervised Hebbian learning are derived from the criterion that the cost (or gain) should be expressed in terms of (generalized) autocorrelation response functions. Our approach also gives an approximate expression for the cooling rate in simulated annealing in terms of heat capacity and energy diffusion. [S1063-651X(96)10409-8]

PACS number(s): 02.70.Lq

An adaptive approach, which has been developed mainly in some branches of engineering, such as control and operational research, is now gathering much attention from statistical physics in connection with Monte Carlo methods and learning in neural network models and so on. By adaptation we do not mean the response to a changing environment or the intrinsic adaptation [1], which has no object function to be maximized. Here it is meant specifically to express slow time variation of system parameters in order to perform calculations more efficiently for our purposes. As examples we only mention temperature control in simulated annealing (SA) [2], some adaptive Monte Carlo methods [3,4], and the Hebbian law in neural network learning [5]. It is to be noted that usually the adaptation process is very slow to ensure that we could arrive at a desired (unknown) attractor with high probability, which gives rise to the possibility of applying linear-response theory (LRT) [6] to some adaptation and learning problems as discussed below.

We take a system with energy $E(\mathbf{x})$ at a phase point $\mathbf{x}$ and introduce for later convenience a density of states $\Omega(E)$ and the entropy $S(E)$ defined by

$$\Omega(E)=\int d\mathbf{x}\ \delta(E-E(\mathbf{x})), \quad S(E)=\ln\Omega(E). \tag{1}$$

First let us consider a general Metropolis Monte Carlo method with the transition probability $W(\mathbf{x}\to\mathbf{x}')$ from the point $\mathbf{x}$ to $\mathbf{x}'$, satisfying the detailed balance condition

$$W(\mathbf{x}\to\mathbf{x}')/W(\mathbf{x}'\to\mathbf{x})=\exp[A(\mathbf{x})-A(\mathbf{x}')]. \tag{2}$$

In the Markovian time series $\{\mathbf{x}_n\}$ generated by $W(\mathbf{x}\to\mathbf{x}')$, the occurrence probability of $\mathbf{x}$ is proportional to $\exp[-A(\mathbf{x})]$. Thus, if we take $A(\mathbf{x})=E(\mathbf{x})/T$, with T the temperature of the system, we have a canonical distribution

$$p_c(\mathbf{x})=\exp[-E(\mathbf{x})/T]/Z_c, \tag{3}$$

for which the energy distribution $\tilde{p}_c(E)$ is obtained from Eq. (1) as

$$\tilde{p}_c(E)\propto\int d\mathbf{x}\ \delta(E-E(\mathbf{x}))e^{-E(\mathbf{x})/T}=e^{S(E)-E/T}\equiv e^{-F(E,T)/T}. \tag{4}$$

When T is small and effects of thermal noise become weak, the time series $\{\mathbf{x}_n\}$ is usually trapped in a local minimum of $E(\mathbf{x})$ for a long time [proportional to $\exp(\Delta E_b/T)$ with ΔE_b the energy barrier], resulting in a nonergodic sampling. This leads to difficulty in dealing with a first-order phase transition because the free energy $F(T,E)$ in Eq. (4) has a double-well structure with a large energy barrier ($\Delta E_b \gg T$) to cross and also an extremely slow cooling rate in the SA [2], which requires the limit $T\to 0$ to attain the global minimum of $E(\mathbf{x})$.

To cope with the disadvantages of the Metropolis sampling (2) with $A=E/T$, a so-called entropy sampling (ES) [7] has been proposed, which enables a direct and adaptive calculation of $S(E)$ as follows. We perform a Monte Carlo simulation with a trial function $A_0(E(\mathbf{x}))$ in Eq. (2) and obtain an energy histogram $h(E)$. Since the equilibrium energy distribution $\tilde{p}_{A_0}(E)$ is proportional to $\exp[S(E)-A_0(E)]$, the region satisfying $S(E)>A_0(E)$ is sampled more often than the region $S(E)<A_0(E)$ and we update $A_0(E)$ by

$$A_1(E)=A_0(E)+\ln h(E) \tag{5}$$

and continue the procedure until we have an E-independent histogram, where our A function is equal to $S(E)$ up to an additive constant [4,8].

To derive and investigate the ES from a microscopic viewpoint, we introduce, corresponding to the Markov process (2), the Langevin dynamics

$$d\mathbf{x}/dt=-\nabla A(E(\mathbf{x}))+\mathbf{f}(t), \tag{6}$$

where the random force $\mathbf{f}(t)=[f_1(t),\ldots,f_N(t)]$ satisfies the fluctuation-dissipation relation

$$\langle f_i(t)f_j(t')\rangle=2\delta_{ij}\delta(t-t'). \tag{7}$$

The Fokker-Planck equation (FPE)

$$\partial p(\mathbf{x},t)/\partial t=\nabla\cdot[p(\mathbf{x},t)\nabla A(E(\mathbf{x}))+\nabla p(\mathbf{x},t)] \tag{8}$$

gives the equilibrium distribution $p_A(\mathbf{x}) \propto \exp[-A(E(\mathbf{x}))]$ as the Markov process Eq. (2) does. Thus it is possible to implement the ES by combining the update algorithm (5) with the Langevin sampling Eq. (6) instead of the Metropolis one, Eq. (2).

The FPE for the energy distribution function $\tilde{p}(E,t)$ $[\equiv \int d\mathbf{x}\ \delta(E-E(\mathbf{x}))p(\mathbf{x},t)]$ is derived by operating $\int d\mathbf{x}\ \delta(E-E(\mathbf{x}))$ on both sides of Eq. (8). Under the assumption that $p(\mathbf{x},t)$ depends on $\mathbf{x}$ through energy $E(\mathbf{x})$, we obtain

$$\partial\tilde{p}(E,t)/\partial t = (\partial/\partial E)[D(E)\{\tilde{p}(E,t)d(A-S)/dE + \partial\tilde{p}(E,t)/\partial E\}] \equiv L\tilde{p}(E,t), \tag{9}$$

where the diffusion constant $D(E)$ in the energy space is defined to be the microcanonical ensemble average of $|\nabla E(\mathbf{x})|^2$,

$$D(E) = \int d\mathbf{x}\ \delta(E-E(\mathbf{x}))|\nabla E(\mathbf{x})|^2 \Big/ \int d\mathbf{x}\ \delta(E-E(\mathbf{x})). \tag{10}$$

It is seen from Eq. (9) that the equilibrium distribution is given by $\tilde{p}_A(E) \propto \exp[-A(E)+S(E)]$.

We now consider the following situation. With a trial function $A(E)=A_0(E)$ we solve the Langevin equation (6) to obtain the sample point $\{\mathbf{x}(t_i)\}$ $(i=1,\ldots,M)$ in the time region $-\tau_0 \leq t \leq 0$. We take τ_0 long so that we can consider that the energy distribution at $t=0$, $\tilde{p}(E,\ t=0)$, calculated from $\{\mathbf{x}(t_i)\}$ $(i=1,\ldots,M)$ is approximately an equilibrium one

$$\tilde{p}(E,t=0) \simeq \exp[S(E)-A_0(E)]/Z_0 \equiv \tilde{p}_{A_0}(E). \tag{11}$$

At $t=0$ we change $A_0(E)$ to

$$A_1(E) = A_0(E) + \delta A(E) \tag{12}$$

and our problem now is how to find the small adaptation $\delta A(E)$ appropriate for our purpose. Since ES aims at the uniform sampling in the energy space [4], we take as an object function the information entropy (difference)

$$M(t) \equiv -\int dE[\tilde{p}(E,t)\ln\tilde{p}(E,t) - \tilde{p}(E,0)\ln\tilde{p}(E,0)] \tag{13}$$

and study how $M(t)$ behaves in response to $\delta A(E)$. Here we regard $\delta A(E)$ as a small perturbation and express the Fokker-Planck operator L, with $A=A_1=A_0+\delta A$ in Eq. (9), as $L_0+\delta L$, where L_0 is the L operator with $A=A_0$ and

$$\delta L \equiv (\partial/\partial E)D(E)\,\delta A'(E), \tag{14}$$

where $\delta A'(E) \equiv d\,\delta A(E)/dE$. Setting $\tilde{p}(E,t) = \tilde{p}_{A_0}(E) + \delta\tilde{p}(E,t)$, we immediately obtain $\delta\tilde{p}(E,t)$ up to a linear response as

$$\delta\tilde{p}(E,t) = \int_0^t \exp(Ls)\,\delta L\,\tilde{p}_{A_0}(E), \tag{15}$$

and from Eq. (13) we have

$$M(t) = \int_0^t ds\langle D(E)\,\delta A'(E)\,\partial\{\ln\tilde{p}_{A_0}\}(E,s)/\partial E\rangle_{\tilde{p}_{A_0}}, \tag{16}$$

where $\langle\ \rangle_{\tilde{p}_{A_0}} \equiv \int dE\ \tilde{p}_{A_0}(E)\cdots$ and

$$\{\ln\tilde{p}_{A_0}\}(E,s) \equiv \exp[L^\dagger s]\ln\tilde{p}_{A_0}(E), \tag{17}$$

with $L^\dagger$ defined to be adjoint to L, i.e., $\int dE\ f(E)Lg(E) = \int dE\ g(E)L^\dagger f(E)$ for arbitrary functions f and g [6(b)]. In order to make $M(t)$ large, or at least a positive quantity, $M(t)$ should be expressed, with some positive measure $p(E)$, as a kind of generalized autocorrelation function

$$M(t) = \int_0^t ds \int dE\ p(E)G(E,s)G(E) \equiv \int ds\langle G(E,s)G(E)\rangle_p \equiv \int ds\ g(s) \tag{18}$$

for which $g(s=0) = \langle G^2(E)\rangle_p > 0$. From the above the first candidate for $\delta A(E)$ is

$$\delta A(E) = \epsilon\ \ln\tilde{p}_{A_0}(E), \tag{19}$$

for which $M(t)$ is expressed as Eq. (18) with $p(E) = D(E)\tilde{p}_{A_0}(E) > 0$ and ϵ is a small constant to control the rate of adaptation. We could also choose δA as defined by $D(E)\,\delta A'(E) = \epsilon d\ \ln\tilde{p}_{A_0}(E)/dE$. In this case δA depends on $D(E)$, of which we have no knowledge in the process of adaptation. We note that Eq. (19) (with $\epsilon=1$) precisely corresponds to the ES update (5). Repeating the procedure (12) and (19), we finally reach the situation where $\ln\tilde{p}_{A_0}(E) = \text{const}$ $[M(t)=0]$ and from Eq. (11) this is equivalent to $A_0(E)=S(E)$ up to an additive constant. This is also consistent with the FPE (9), which gives a uniform distribution when $A=S$.

As an example of the ES we consider a system [9] of N $(=10)$ coupled Duffing oscillators

$$E(\mathbf{x}) = \sum_{i=1}^{N}[x_i^2/2 + x_i^4/4] - \{2B/(N-1)(N-2)\} \times \sum_{i<j<k} x_i x_j x_k, \tag{20}$$

which approximately models mode dynamics for the liquid-solid phase transition with x_i denoting Fourier amplitude of a density wave for a crystalline solids. In Fig. 1 we show how $A(E)$ converges to $S(E)$ under the adaptation (5) for the time increment of 10^5 Monte Carlo steps, starting from $A(E)=0$ [$B=3$ in Eq. (20)]. We see that it takes long time before $A(E)$ gets some weight in the region $E<0$ and the origin for this slow penetration of the energy distribution across $E=0$ can be traced to $D(E)$ [Eq. (10)], shown in Fig. 2, which is small near $E=0$. From Fig. 1 a double tangent can be drawn for $B=3$, indicating that a first-order transition at $T=T_c$ occurs that is not sharp. Similiar nonsharp transition is also observed in the model of protein folding [10].

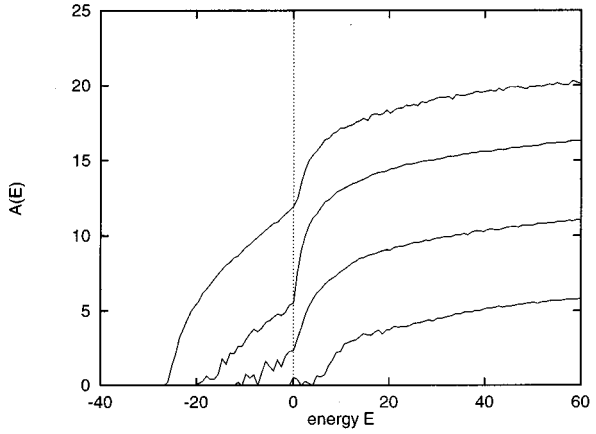

FIG. 1. Convergence of $A(E)$ to $S(E)$ with 10^5 Monte Carlo steps for each adaptation ($B=3$). Initially $A(E)$ is set equal to 0.

From the relation $1/T=(\partial S/\partial E)$ we have $T_c \simeq 2.5$. Finally, in Fig. 3 we show the order-parameter distribution $p(X)(X \equiv \Sigma x_i/N)$ for $T=5.0$, 2.5, and 1.7 ($B=3$). This is obtained by first calculating $S(E,X) \equiv \int d\mathbf{x}\, \delta(E-E(\mathbf{x}))\delta(X-X(x))$ by the (generalized) ES and then performing E integration $p(X) \propto \int dE \exp[-E/T+S(E,X)]$. We observe a transition of $\langle X \rangle$ from $\langle X \rangle \simeq 0$ to $\langle X \rangle \simeq 2.5$ as T decreases from 5.0 to 1.7. For $B<2.7$, $p(X)$ has a one-peak structure for all temperature and we have no phase transition.

Before proceeding to learning in a neural network, we study the problem of cooling rate in the SA [2] based on LRT and the Langevin model (6). Here it is noted that $A(E)$ is replaced by E itself and the relation (7) becomes $\langle f_i(t) f_j(t') \rangle = 2T\delta_{ij}\delta(t-t')$. The FPE now takes the form

$$\partial p(\mathbf{x},t)/\partial t = \boldsymbol{\nabla} \cdot [p(\mathbf{x},t)\boldsymbol{\nabla} E(\mathbf{x}) + T\boldsymbol{\nabla} p(\mathbf{x},t)] \equiv L_{\mathrm{FP}} p, \tag{21}$$

with the equilibrium distribution given by Eq. (3). The FPE for the energy distribution is readily derived from Eq. (21), as before, to have

$$\partial \tilde{p}(E,t)/\partial t = (\partial/\partial E)[D(E)\{\tilde{p}(E,t)\partial F(E,T)/\partial E + T\partial \tilde{p}(E,t)/\partial E\}] \equiv L\tilde{p}, \tag{22}$$

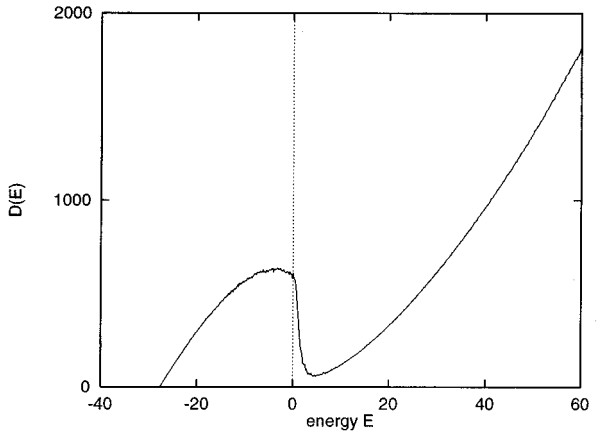

FIG. 2. Diffusion constant $D(E)$ [Eq. (10)] obtained from the ES simulation ($B=3$).

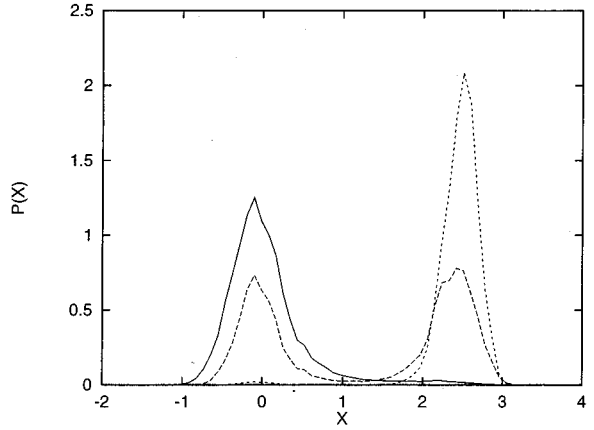

FIG. 3. Probability distribution function $p(X)$ for $T=5.0$ (full curve), $T=2.5$ (dashed curve), and $T=1.7$ ($B=3$).

with $F(E,T)$ and $D(E)$ defined by Eqs. (4) and (10), respectively.

The situation for the SA is as follows. For time $t<0$ our system is assumed to be in equilibrium at temperature T; thus $\tilde{p}(E,\, t=0) \simeq \exp[S(E)-E/T]/Z_c$. At $t=0$ we decrease T by δT, $T(t>0)=T-\delta T$, and we are interested in how the average of energy $\langle E(t) \rangle$ behaves in response to annealing. Since we assume that δT is small we can apply LRT to calculate $\langle E(t) \rangle \equiv \int dE\ \tilde{p}(E,t)E$. Precisely following the argument that led to Eq. (16), we readily obtain the energy drop

$$|\langle E(t) \rangle - \langle E(0) \rangle| = (\delta T/T) \int_0^t ds \langle D(E)\partial E(E,s)/\partial E \rangle_{\tilde{p}_c}, \tag{23}$$

where $E(E,s) \equiv \exp(L^{\dagger}s)E$, with $L^{\dagger}$ denoting the adjoint operator of L, Eq. (22). It is interesting to note that since $E(E,0)=E$ we can express the integrand in Eq. (23) in the form of a (generalized) autocorrelation function $\langle \partial E(E,s)/\partial E \partial E(E,0)/\partial E \rangle_{D(E)\tilde{p}_c}$. Actually, with some manipulation on Eq. (23) or directly applying LRT to Eq. (21), we have an equivalent expression

$$|\langle E(t) \rangle - \langle E(0) \rangle| = (\delta T/T) \int_0^t ds \langle \boldsymbol{\nabla} E(\mathbf{x},s) \cdot \boldsymbol{\nabla} E(\mathbf{x},0) \rangle_{p_c}, \tag{23'}$$

where $E(\mathbf{x},s) \equiv \exp(L_{\mathrm{FP}}^{\dagger}s)E(\mathbf{x})$ with $L_{\mathrm{FP}}^{\dagger}$ an adjoint operator of L_{FP} in Eq. (21). Denoting by τ_R the relaxation time of the correlation function and noting that $\partial E(E,\, s=0)/\partial E=1$, we can estimate Eq. (23) approximately as

$$|\langle E(t>\tau_R) \rangle - \langle E(0) \rangle| \simeq (\delta T/T) \langle D(E) \rangle_{\tilde{p}_c} \tau_R. \tag{24}$$

On the other hand, it is easy to calculate the average energy at $T-\delta T$, $\langle E \rangle_{T-\delta T}$ to obtain

$$|\langle E \rangle_{T-\delta T} - \langle E \rangle_T| = \delta T \langle (E-\langle E \rangle_T)^2 \rangle_T / T^2 = \delta T C(T), \tag{25}$$

with $C(T)$ the specific heat of the system. Equations (24) and (25) yield an expression for the cooling rate R_c [11],

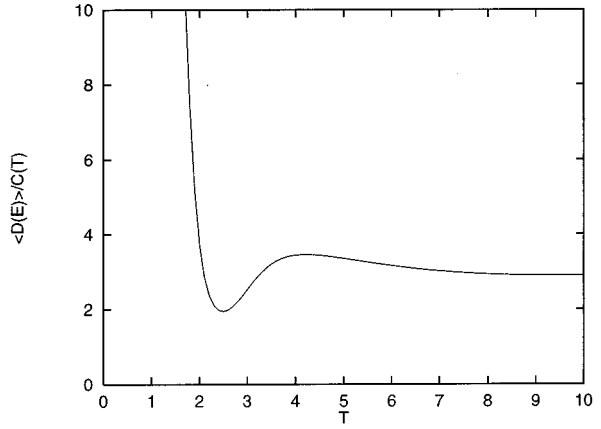

FIG. 4. $\langle D(E)\rangle_{\tilde{p}_c}/C(T)$ as a function of temperature ($B=3$).

$$R_c(T)\equiv \delta T/\tau_R=\delta T\langle D(E)\rangle_{\tilde{p}_c}/[TC(T)]. \tag{26}$$

Equation (26) states that when C becomes large, slow cooling is required, as noticed by Kirkpatrick, Gelatt, and Vecchi [2]. Also careful cooling is necessary, as noted in connection with Figs. 1 and 2, when $\langle D(E)\rangle_{\tilde{p}_c}$ becomes small because it takes a long time for an excursion in the energy space. From Fig. 4, which depicts $\langle D(E)\rangle_{\tilde{p}_c}/C(T)$ for the model (20) ($B=3$) as a function of T, it is seen that the rate has a minimum $R_{c,\min}$ around the transition point $T=T_c$. In order to avoid trapping in an nonequilibrium state or a glass transition we must keep the cooling rate smaller than that given by Eq. (26). It is to be noted that Eq. (26) gives a general but rather rough estimate of the cooling rate because in actual SA processes the Kramers time $\exp(\Delta E_b/T)$ plays an important role, as stated below the line of Eq. (4).

Finally, let us consider a network model consisting of N formal neurons (Ising spins), with the Hamiltonian

$$E(\mathbf{s})=-\tfrac{1}{2}\sum_{i,j\ (i\neq j)} J_{ij}s_is_j. \tag{27}$$

Glauber dynamics of the system is described by the master equation [12]

$$\partial p(s_1,\ldots,s_N;t)/\partial t=-\sum_i\sum_{s_i'=\pm 1} s_is_i' f(-s_i'|h_i) \times p(s_1,\ldots,s_i',\ldots,s_N;t)\equiv L_Gp, \tag{28}$$

where the transition probability of spin i from $-s_i$ to s_i under the field $h_i=\Sigma_{j\ (\neq i)}J_{ij}s_j$ is defined to be

$$f(s_i|h_i)=\exp(h_is_i/T)/[\exp(h_i/t)+\exp(-h_i/T)]. \tag{29}$$

The (canonical) equilibrium distribution function p_c is given by

$$p_c(\mathbf{s})=\exp\left(\sum_i s_ih_i/(2T)\right)\Big/ Z_c. \tag{30}$$

Now the situation we are interested in is as follows. At $t=0$ the system is in equilibrium, as described by Eq. (30) with some interaction J_{ij}. In order to embed a pattern $\{s_i\}=\{\xi_i\}$ $(i=1,\ldots,N)$ in the network we change h_i to $h_i+\delta h_i$ at $t=0$ and try to find the desirable adaptation δh_i based on LRT. As in Eq. (15) we have

$$p(\mathbf{s};t)=p_c(\mathbf{s})+\int_0^t ds\ \exp[L_Gs]\delta L_Gp_c(\mathbf{s}) \tag{31}$$

and δL_G is readily seen from Eq. (28) and

$$f(s_i|h_i+\delta h_i)=f(s_i|h_i)[1+(\delta h_i/T)\{s_i-\tanh(h_i/T)\}] \tag{32}$$

to be given by

$$\delta L_Gp_c(\mathbf{s})=(2/T)\sum_i\ \delta h_is_iQ(\mathbf{s})p_c(\mathbf{s}), \tag{33}$$

where $Q(\mathbf{s})\equiv\{\exp(-h_is_i/T)/[\exp(h_i/T)+\exp(-h_i/T)]\}$ is positive definite.

Intuitively one may take the overlap [12]

$$m(t)=\sum_{\mathbf{s}}\ p(\mathbf{s};t)\mathbf{s}\cdot\boldsymbol{\xi}/N\equiv\langle\mathbf{s}(t)\rangle\cdot\boldsymbol{\xi}/N \tag{34}$$

as the appropriate object function to be maximized. The response to δh_i is

$$\delta m(t)=(2/NT)\int_0^t\langle\boldsymbol{\xi}\cdot\mathbf{s}(s)\,\delta\mathbf{h}\cdot s(0)\rangle_{Qp_c}, \tag{35}$$

with $\mathbf{s}(s)=\exp[L_G^\dagger s]\mathbf{s}$ as usual. From Eq. (35) we immediately notice that by setting

$$\delta\mathbf{h}=\epsilon\boldsymbol{\xi}\quad\text{or}\quad\delta h_i=\epsilon\xi_i\ \ (i=1,\ldots,N), \tag{36}$$

we have a generalized autocorrelation function expression for $\delta m(t)$. The choice (36) is nothing but an external (constant) field along the pattern. As another candidate for the object function we can take

$$p(\mathbf{s}=\boldsymbol{\xi};t)=\langle\delta(\mathbf{s}-\boldsymbol{\xi})\rangle. \tag{37}$$

Equation (37) means that one should increase the probability that the system takes the configuration $\mathbf{s}=\boldsymbol{\xi}$. This time let us employ the adaptation $J_{ij}\to J_{ij}+\delta J_{ij}$ or

$$\delta h_i=\sum_{j\ (\neq i)}\delta J_{ij}s_j \tag{38}$$

and from Eqs. (37) and (33) we have

$$\delta p(\mathbf{s}=\boldsymbol{\xi};t)=(2/T)\int_0^t ds\left\langle\{\exp(L_G^\dagger s)\,\delta(\mathbf{s}-\boldsymbol{\xi})\}\sum_{i,j}\ \delta J_{ij}s_is_j\right\rangle_{Qp_c}. \tag{39}$$

Since in the integrand $\mathbf{s}=\boldsymbol{\xi}$ at $s=0$ we are led to the Hebbian rule [12]

$$\delta J_{ij} = \epsilon \xi_i \xi_j \quad (\epsilon > 0) \tag{40}$$

from the condition that the integrand is positive at $s=0$.

In this paper some problems related to the ES [4], the SA [2], and learning in a neural network [12] were studied based on LRT, which has been traditionary used to study the response of physical systems to external fields [6]. It is hoped that LRT could shed some light on wider problems in learning and information processing, which now gather much interest from many branches of natural science.

[1] N. H. Packard, in *Artificial Life, SFI Studies in the Science of Complexity*, edited by C. Laugton (Addison-Wesley, Reading, MA, 1988).

[2] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).

[3] B. A. Berg and T. Neuhans, Phys. Rev. Lett. **68**, 9 (1992).

[4] J. Lee, Phys. Rev. Lett. **71**, 211 (1993).

[5] J. Herz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Reading, MA, 1991).

[6] (a) R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957); (b) T. Munakata, Phys. Rev. E **50**, 2351 (1994).

[7] That the ES [4] and the multicanonical sampling [3] are equivalent is shown by B. A. Berg, U. H. E. Hansmann, and Y. Okamoto, J. Phys. Chem. **99**, 2236 (1995). In this paper we take the ES approach for convenience of presentation.

[8] J. Lee and M. Y. Choi, Phys. Rev. E **50**, R651 (1994).

[9] T. Munakata, J. Phys. Soc. Jpn. **60**, 2800 (1991); S. A. Alexander and J. McTague, Phys. Rev. Lett. **41**, 702 (1978).

[10] Ming-Hong Hao and Horald A. Scheraga, J. Chem. Phys. **102**, 1334 (1995).

[11] δT must be small enough to ensure the applicability of the LRT.

[12] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, England, 1989).