STOCHASTIC DYNAMICS AND REPLICA SYMMETRY BREAKING OF NEUROSYSTEMS

Yoshitake Yamazaki[a], Su-Jin Kim[a,b] and Moyuru Ochiai

a.Computer Science and Systems Engineering, Kyushu Institute of Technology, Iizuka, 820, Japan

b.Electronics, National Fisheries University of Pusan, Pusan, 608-737, Korea

c.Electronics, North-shore College, Atsugi, 243, Japan

Abstract

Subject on neural networks (NNW) is very important, which is strongly concerned with nonequilibrium (NEQ) and/or equilibrium (EQ) statistical physics in random media, e.g. crystal growth, glasses, polymers and vortices. To establish foundations of such physics, stochastic dynamics and replica symmetry breaking (RSB) of generalized neurosystems are discussed using supersymmetry (SUSY) fields. The following items are concluded; (1) the valid range of replica symmetric (RS) solution, in the RSB-regime, (2) the validity of the hierarchical structure (HS), (3) the NEQ-behavior (e.g. aging) to a metastable state and (4) the partial annealing (PA) properties.

1 INTRODUCTION

Aims of study on neural networks are to find and to construct some of the best networks with human functions, e.g. learning, association and storing. In a generalized model of NNWs we establish stochastic dynamics using SUSY-fields and compare the results obtained with the replica solutions. Based on the SUSY-stochastic dynamics we derive a true behavior of the system and the regime where the RS-theory gives a valid solution to the system. We also discuss RSB-behavior, in addition to the PA-properties.

2 Neural Network Model

Consider a generalized soft m-neuron model ($m \geq 2$) specified with the cost function

$$E = -\Sigma_{1\leq i_1<\cdots<i_m\leq N} J_{i1\cdots im} s_{i1}\cdots s_{im} - h\Sigma_i s_i + h\{s\} \quad \text{with the interactions among}$$

$$\text{neurons like the Hopfield network } J_{i1\cdots im} = \frac{1}{N}\Sigma_{\mu=1}^{p}\xi_{i1}^{\mu}\cdots\xi_{im}^{\mu} \quad \text{for } i_1,\cdots,i_m \text{ different neuron}$$

and the neuron value probability $h\{s\}$ for the spherical model or with multiple peaks ($s = 0, \pm1, \pm2, \cdots$). The system learns p patterns $\{\xi^{\mu}\}(\mu = 1,\cdots,p;\ \xi = \pm1)$. Study the neuron states $\{s\}$ at time t in the recalling process after one of the teacher patterns was randomly chosen.

3 SUSY Stochastic Dynamics

Due to the stochastic Langevin dynamics, probability $P(s,t)$ of finding the system in a point $\{s\}$ of phase space time-evolves according to the Fokker-Planck (FP) equation which is described as

$$Z = \int D[\phi]e^{-s_k - s_p},\ S_k \equiv -\frac{1}{2}\int d\theta d\bar{\theta}dt\ \Sigma_i\phi_i D^{(2)}\phi_i,\ S_p \equiv \int d\theta d\bar{\theta}dt E(\phi), \tag{1}$$

using fermionic variables $\{\eta,\ \bar{\eta}\}$, anticommuting Grassmann variables $\{\bar{\theta},\ \theta\}$ and superfields $\phi_i \equiv s_i + \bar{\theta}\eta_i + \bar{\eta}_i\theta + p_i\bar{\theta}\theta$, $\{p_i \equiv -i\partial/\partial s_i,\ a \equiv (t_a, \bar{\theta}_a, \theta_a)\}$ with $D^{(2)} \equiv 2\partial^2/\partial\theta\partial\bar{\theta} + 2\theta\partial^2/\partial\bar{\theta}\partial t - \partial/\partial t$. All correlation and response functions $\{C(t_a,t_b) = \Sigma_i\langle s_i(t_a)s_i(t_b)\rangle/N,\ G(t_a,t_b) = \Sigma_i\langle\delta s_i(t_a)/\delta h(t_b)\rangle/N\}$ are expressed with the SUSY-correlation function

$$\langle Q(a,b)\rangle = \Sigma_i\langle\phi_i(a)\phi_i(b)\rangle/N = C(t_a,t_b) + (\bar{\theta}_b - \bar{\theta}_a)[\theta_b G(t_a,t_b) + \theta_a G(t_b,t_a)]. \tag{2}$$

The auto-correlation is associated with the order parameter (OP) q of a metastable state [which is concerned with a quantity like the width of its valley]. To study the NEQ-behavior and the HS-validity we adopt the following change of time-variable in the correlation and response functions:

$$C(t_1,t_2) = b(C(t_1,t_2)) = b(C(t_1,t_3),C(t_3,t_2)),\ G(t_1,t_2) = \theta(t_1-t_2)a(C(t_1,t_2))\partial C(t_1,t_2)/\partial t_2 \tag{3}$$

where the functions b, a are determined from the dynamics of the system.

4 Results

Based on the SUSY-stochastic dynamics we studied the cases of $m = 2$ and $m = \lambda + 1$[1, 2, 3] but here discuss the former case. The method is as follows: (1) Averaging over both the recalling pattern ($\mu = 1$) and nonrecalling ones ($\mu = 2, 3, \cdots, p$) we derive the leading expression for the SUSY-partition function in the neuron number N. (2) From the behavior of the system in the limit of large neuron-number ($N \to \infty$) (i.e. at a single saddle point), the expressions for the OP are derived. (3) By means of the fluctuation-dissipation theorem (FDT) we solve the eigenvalues of the quadratic fluctuations around the saddle point.

The results obtained in the RS-regime are as follows: The expressions for all OPs of the system coincide with those derived in terms of the replica method, but the regime where the replica solutions hold valid is limited only in the regime where the FDT holds valid.

In the RSB-regime the following characteristic results were derived.

(A) Validity of the HS The HS should be specified with a set of OPs $\{q\}$ which is derived from the correlation function (3). This expression corresponds to the renormalization equation for $\{q\}$ expressed by the renormalization operator b. A series of the fixed points obtained there satisfies $C(t_1^*, t_2^*) = \min\ \{t_1^*, t_2^*\}$ (i.e. $q(x) = q_i$ for $m_i \leq x \leq m_{i+1}$), i.e. the HS is constructed with the set of the fixed-point OPs $\{q_i\}$. This structure is considered to be considerably valid.

(B) Nonequilibrium behavior In case of the paraphase the system displays a behavior ergodically decaying towards the equilibrium state, i.e. $a(z) = 1$ for $q < z < 1$ in (3). On the other hand, in case of metastable phases the system tends to a nonergodic behavior approaching a finite metastable state, i.e. $a(z) \neq 1$ for $z = q$, which changes depending on the step of the HS, the temperature and the external field strength. As the nonergodic properties consider the aging properties in a case of any one valley hierarchically splitting into two valleys. The system has a function of the long-termed memory and draws the hysteresis loop of its history-dependence. From the change of the shape we can know the details of the HS and the step working in the HS.

(C) Partial annealing To make clear the annealing problems of the system, we denote temperature for the environment of neurons (interactions acting among them) by T (T', respectively). Then the replica index is related as $n = (T/T')^{1-c}$ (c:thermal coupling rate). It is shown that the regions of (para, glass, ferro, superferro and RSB) phases and their phase boundaries change, depending on the temperature ratio, the step of the HS and the storage capacity. Their dependences of dynamical behavior in these phases were made clear.

5 Concluding remarks

We could establish the SUSY-stochastic dynamics for the generalized NNWs. In the RS-regime we could show that the replica solutions gave the valid results in the regions where the FDT was satisfied. Furthermore, we could make clear the characteristic properties of the system in the RSB-regime.

As the properties of the NNW-system the most important point is that the nonergodic features of the hysteresis loops can be utilized for the intelligent control of the neurosystems. On the other hand, it was found that NEQ-systems (such as the crystal growth, glasses and polymers) could be described with the OPs in the HS, their time-evolutions could be treated by means of the SUSY-stochastic dynamics, and their behavior was characterized by the hysteresis loops.

References

[1] Y.Yamazaki, M.Ochiai and A.Holz Proc. 3rd Int. Conf. on Fuzzy Logic, Neural Nets & Soft Comp. 115 (1994).

[2] Y.Yamazaki, M.Ochiai and A.Holz Proc. of KFMS Conf. '94, 85 (1994).

[3] Y.Yamazaki The 19th IUPAP Int. Conf. of Statistical Physics, 186 (1995).