Information dynamics of neural networks with the aid of supersymmetry fields as the microscopic thermal flow

2000 J. Phys. A: Math. Gen. 33 8389

(http://iopscience.iop.org/0305-4470/33/47/306)

J. Phys. A: Math. Gen. **33** (2000) 8389–8407. Printed in the UK PII: S0305-4470(00)12890-2

Information dynamics of neural networks with the aid of supersymmetry fields as the microscopic thermal flow

Yoshitake Yamazaki

Institute of Fluid Science, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

Received 28 March 2000, in final form 15 August 2000

Abstract. Information dynamics is discussed from the point of view of microscopic thermal flow. In the fields of learning, association, storage and so on, the concepts of neural networks (NNWs) have been widely used. Their neurodynamics have been investigated as a stochastic process of an infinite neuron system, using the replica method. This approach includes unsettled points; regimes where replica symmetry (RS) solutions and replica symmetry breaking (RSB) solutions are valid, low-dimensional behaviour regimes, and so forth.

First, to make them clear, using supersymmetry (SUSY) fields the dynamics of NNWs are investigated for a family of NNWs interacting among m neurons. The dynamics of the system are supposed to be specified according to the Langevin dynamics with Gaussian white noise (i.e. a random influence from the surroundings) under an environmental parameter β (such as the inverse temperature). The results obtained without ambiguity are as follows: the RS solutions are valid in the regime where our solutions satisfy the fluctuation–dissipation theorem (FDT), while the RSB solutions appear in the SUSY-breaking regime. As a function of the environmental parameter, the system displays transitions from usual (ergodic) phases to phases with broken ergodicity.

Secondly, the information dynamics of NNW is derived as the microscopic thermal flow.

1. Introduction

The aims of the study of neural networks (NNWs) are to find and to construct some of the best networks with human functions, e.g. learning, association and storage on one side and to derive the relation of the information dynamics to the microscopic thermal flow on the other side. Models of NNWs have been generalized as far as possible with a set of external parameters. Methods for analysing them have also been improved with various methods, e.g. the replica and Monte Carlo methods for studying their stochastic dynamics [1–11]. Characteristics of the system can be specified from various points of view, but here we restrict to a Q (quenched) case; randomly choosing one pattern out of a set of patterns $\{\xi\}$ we average neural quantities over the neuron state variables $\{s\}$ under the randomly chosen pattern, i.e. we regard the neuron pattern (state) variables as the slow (fast) variables of the system (respectively). In this Q-case the replica method is one of the best available methods. The average cost function (free energy) is written as

$$\langle\langle \ln Z \rangle\rangle = \lim_{n\to 0}[\langle\langle Z^n \rangle\rangle - 1]/n \tag{1.1}$$

where Z^n denotes the n replicas of the system (a partition function), whose variables $\{s_i\}$ are replicated as $\{s_i^\alpha\}$ ($\alpha = 1, 2, \ldots, n$). From the mathematical point of view, the expression (1.1) is identity for Z positive, otherwise it is wrong. Physically, the range of validity of (1.1) should be determined and restricted from the dynamical instability of the system. The boundary between the replica symmetry (RS) and the replica symmetry breaking (RSB) states is called the AT line, and its systematic determination is done by the system instability within the framework of the replica theory. It needs to be shown that the results derived in this theory coincide with those evaluated by another theoretical formalism, on one hand, and also that their behaviours are related to a traditional relation, such as the microscopic thermal flow (heat transfer equation), on the other hand. These subjects are investigated in this paper.

The first subject will be made clear in sections 2–4, using a powerful and systematic method, the supersymmetry (SUSY) field method for the evaluation of the characteristics of the NNWs [7]. In terms of the SUSY fields we study the stochastic dynamics of a family of generalized NNWs interacting among m ($\geqslant 2$) neurons in the Q-phase. The SUSY theory shows that the RS theory gives valid results only in the regime where the fluctuation–dissipation theorem (FDT) holds.

In section 2 the model NNWs are specified—so-called generalized soft m-neuron models [4–6]. There a neuron state $s_i(t+1)$ in the ith neuron at time $t+1$ is determined using the sum of products among different m-multiplicative sets of patterns and neuron states at time t over all different combinations of neurons. The neuron state is assumed to take continuous values. In the case of a neuron state with two values (± 1) we introduce a neuron state probability with sharp double peaks and obtain the same results as in the discrete neuron case.

In section 3 a method of SUSY stochastic dynamics is developed. In order to avoid the replica method and to take the random average systematically, we introduce fermions and SUSY fields [3, 7]. The generalized partition function of the SUSY stochastic-dynamic system and the important fluctuation–dissipation theorem and causality relation are derived. It is shown that the two-neuron correlation function in the superspace contains the two-neuron correlation and response functions of the neuron system.

In section 4 the results for the cases of $m = 2$ and $m > 2$ are derived and discussed in subsections 4.1 and 4.2, respectively. The first and the second cases have already been studied in [2, 4] and [2, 5, 6], respectively, in terms of the replica method. It is shown that the SUSY stochastic dynamics yield the same results as those obtained using the replica method and that the stability of the system, the RSB regime, and so on can be easily discussed without ambiguity.

Secondly, the second subject is investigated by taking into account the fact that fluctuations of neuron states are subject to the central limit theorem in systems with large numbers of neurons $N \gg 1$. From this we derive the relation of the information dynamics (discussed in sections 2–4) to the microscopic thermal flow in section 5. The essential contributions of the information dynamics are concluded to be derived from the concept of the microscopic thermal flow. The informational flows correspond to it. (a) The pure states are associated with the metastable multivalleys of the system potential. (b) The overlapping parameters correspond to the overlap between their valleys. (c) The distribution of barriers and valleys in the potential, i.e. the distribution of the local temperatures in space, strongly affects the system behaviour. All behaviours are understood as the microscopic thermal flow.

Finally, concluding remarks are summarized in section 6.

2. Model neural networks

Various types of NNW models have been constructed and their properties have been investigated. Many of them are widely used in the fields of artificial intelligence, computer science and engineering. In order to be able to discuss properties of NNWs in various structures we consider the generalized soft m-neuron models ($m \geqslant 2$) specified by the following cost function:

$$E = -\sum_{1\leqslant i1\ldots im\leqslant N}\sum J_{i1\ldots im}s_{i1\ldots}s_{im} - h\sum_{1} s_i + h\{s\} \qquad (ik = i_k \text{ for } k = 1,\ldots,m) \tag{2.1}$$

where the interactions acting among neurons are assumed to be expressed as for the Hopfield network,

$$J_{i1\ldots im} = 1/N\sum_{\mu=1}^{p}\xi_{i1}^{\mu}\ldots\xi_{im}^{\mu} \qquad (i_l,\ldots,i_m \text{ different neuron}). \tag{2.2}$$

Note that the successive subscripts, e.g. ik denote i_k $(k = 1,\ldots,m)$ in this paper. When we consider a layered-structural NNW, the sum over the combination of $i_l,\ldots,i_m$ means that we choose all possible different sets $\{i_l,\ldots,i_m\}$ out of neurons on each layer. In the case of mutually connective NNWs it means that all possible, different sets $\{i_l,\ldots,i_m\}$ are selected from the total number of neurons N of the system.

The NNW system learns p patterns, denoted by $\{\xi^{\mu}\}$ $(\mu = 1, 2,\ldots,p)$ (whose component ξ^{μ} is given with discrete values $\pm l$). As a teacher pattern, one pattern is chosen randomly out of the pattern set $\{\xi\}$ and we consider the neuron signals (states) $\{s\}$ at time t as the recalling process.

The neuron state variables $\{s\}$ are assumed to be of continuous variables. (Even if they are of discrete variables, the obtained results do not change within the approximation of quadratic form in $\{s\}$.) To do so, the neuron value probability $h\{s\}$ is introduced for the spherical model:

$$h\{s\} = \delta\left(\sum_i s_i^2 - N\right). \tag{2.3}$$

Here the spherical probability is chosen only for simplicity of explanation of the SUSY dynamics, but our method is also applicable to various cases, e.g. probabilities with multiple peaks (which mean that $s = 0, \pm1, \pm2, \ldots$) such as

$$h\{s\} = \sum_i [1/2r_0 s_i^2 + us_i^4] \tag{2.4}$$

for double peaks.

3. Method of SUSY stochastic dynamics

Neuron systems such as the brain are constructed from enormous number of neurons. The neurodynamics of such systems can be studied within the framework of the Langevin dynamics with an environmental (inverse temperature) parameter β

$$\Gamma_0^{-1}\,\mathrm{d}s_i/\mathrm{d}t = -\partial(\beta E)/\partial s_i + \nu_i(t) \tag{3.1}$$

where the Gaussian random (noise) variables $\{\nu\}$ with zero mean $(\langle\nu(t)\rangle = 0)$ and variances

$$\langle\nu_i(t)\,\nu_i(t')\rangle = 2/\Gamma_0\delta_{ij}\delta(t-t') \tag{3.2}$$

are introduced. The average $\langle\cdots\rangle$ is taken over the Gaussian distribution characterized with a parameter Γ_0.

The probability of finding the system in a point $\{s\}$ of phase space, $P(s,t)$, evolves in time according to the Fokker–Planck (FP) equation:

$$\partial P/\partial t = \Gamma_0\sum_i \partial/\partial s_i[\partial/\partial s_i + \beta\partial E/\partial s_i]P. \tag{3.3}$$

We rewrite this FP equation, using the SUSY fields (A.13) in appendix A, as

$$
\begin{aligned}
Z &= \int D[\phi]\mathrm{e}^{-Sk-Sp} \\
S_k &\equiv -\tfrac{1}{2}\int \mathrm{d}\theta\,\mathrm{d}\bar{\theta}\,\mathrm{d}t \sum_i \phi_i D^{(2)}\phi_i \\
S_p &\equiv \int \mathrm{d}\theta\,\mathrm{d}\bar{\theta}\,\mathrm{d}t\; E(\phi)
\end{aligned}
\tag{3.4}
$$

where $Sk \equiv S_k$ and $Sp \equiv S_p$ denote the kinetic and potential terms, respectively.

From the information of the SUSY correlation function $\langle Q(a,b)\rangle$ in (A.17), all properties of the correlation and response functions of neuron systems such as $\langle s_i(t_a)\, s_i(t_b)\rangle$, $\langle s_i(t_a)\; p_i(t_b)\rangle$, etc are derived.

4. Results: qualities of NNWs

4.1. Case of a synaptic junction of order $m = 2$

We consider the neural system of order $m = 2$ in expression (2.1). An equilibrium state $\{s\}$ of the system is characterized by the mean overlap with the set of stored patterns $\{\xi^\mu\}$

$$m^\mu = 1/N \sum_i \xi_i^\mu \langle s_i\rangle. \tag{4.1}$$

As only k components of the overlap vector $\{m^\mu\}$ are non-vanishing in the limit $N \to \infty$, we consider the case $k = 1$, i.e. the pattern $\mu = 1$ ($\mu = 2, \ldots, p$) corresponds to the recalling (non-recalling) pattern (respectively). This means $m^1 \sim 1$ and $m^\mu \sim 0$ ($\mu = 2, \ldots, p$). Besides the overlap parameters (OPs) $\{m^\mu\}$, two further OPs are concerned;

$$r = N/p \sum_{\mu>k}^{p} \langle\langle (m^\mu)^2 \rangle\rangle \quad (\text{here } k = 1) \qquad p = \left\langle\!\!\left\langle 1/N \sum_i \langle s_i\rangle^2 \right\rangle\!\!\right\rangle. \tag{4.2}$$

Applying the SUSY stochastic dynamics to this neural system, we obtained fundamentally important properties of the NNW system within the framework of the saddle-point approximation in the limit $N \to \infty$ (see appendix B), which coincide with those derived in terms of the replica method [4].

Let us summarize the important results obtained.

4.1.1. Discrete limit of neuron states $s = \pm 1$. The set of coupled equations for OPs $\{m, q, r\}$ are given from (B.18) and (B.19) as

$$
\begin{aligned}
m &= \int Dz\, \mathrm{th}\, \beta(z(\alpha r)^{1/2} + m) \\
q &= \int Dz\, \mathrm{th}^2\, \beta(z(\alpha r)^{1/2} + m) \\
r &= q/[1 - \beta(1-q)]^2
\end{aligned}
\tag{4.3}
$$

where Dz is defined below (B.18).

The three eigenvalues of the quadratic fluctuations around the saddle point are of the non-degenerate eigenvalue Λ_1, and twofold degenerate one Λ_2;

$$\Lambda_1 = \beta[1 - \beta(1-q)]^{-2} \qquad \Lambda_2 = \Lambda_1 - 2\beta^2 q[1 - \beta(1-q)]^{-3} \tag{4.4}$$

which are given in (B.24). The positive, maximum eigenvalue stabilizes the symmetric system, while negative eigenvalues make the system inverse (i.e. symmetry breaking).

Furthermore, the eigenvalues of the stability matrix for the symmetric solutions (which correspond to the RS solutions) are obtained in (B.28) and (B.29) as

$$\lambda_{\pm} = \alpha\beta[-(u+v) \pm \{(u-v)^2 + 4\}^{1/2}] \tag{4.5}$$

where

$$u \equiv \alpha\beta^2 \langle\!\langle (1 - \langle s \rangle^2)^2 \rangle\!\rangle \qquad v \equiv [1 - \beta(1-q)]^{-2}. \tag{4.6}$$

For $\beta < 1/(1-q)$, the eigenvalue λ_- is negative, while $\lambda_+ > 0$ if $uv < 1$, otherwise $\lambda_+ < 0$; that is, $uv = 1$ gives the boundary of instability to the symmetry breaking regime. If the temperature T is higher than the glass-phase temperature T_g, the symmetric solutions are stable because $uv < 1$.

From the functional point of view of the NNWs, the fundamental properties are summarized as follows. It is considered that the environment parameter β expresses the influence of the environment of the NNW or the ability of the neurons. The characteristics of the system are specified according to values of β; the limit $\beta \to \infty$ corresponds to the ideal perfect case and $\beta \to 0$ to the opposite case. From the numerical computation for three coupled equations in (4.3) the following conclusions are summarized.

(a) The maximum storage capacity $\alpha_c \cong 0.138$ exists in the limit $\beta \to \infty$, i.e. $\alpha < \alpha_c$ for any β and in such small-α regimes the retrieval OP m has larger values than $m \cong 0.973$ and behaves like

$$m = 1 - \exp(-1/2\alpha). \tag{4.7}$$

(b) As the average percentage of errors N_e/N is defined as $(1-m)/2$, the NNW can effectively retrieve the teacher pattern in the regime below α_c, while it reaches the unconscious state (50% errors) near α_c or near T_M:

$$\begin{aligned} N_e/N &\cong (\alpha/2\pi)^{1/2} \exp[-1/2\alpha] \qquad &(\beta \to \infty, \alpha \text{ small}) \\ T_M &\cong (\alpha_c - \alpha)/c_o\alpha_c \qquad &(c_o \cong 0.18) \end{aligned} \tag{4.8}$$

which are derived using the asymptotic behaviour of the error function.

(c) The transition temperature T_g from the disordered neuron phase to the neuron glass (NG) phase is expressed as

$$T_g(\alpha) \cong 1 + \alpha^{1/2} \tag{4.9}$$

from the leading term for q,

$$q \cong \beta^2\alpha q/(1-\beta)^2. \tag{4.10}$$

(d) Below $T_g(\alpha)$, the line $T_M(\alpha)$ appears, below which the retrieval solutions become locally stable, i.e. they have a macroscopic overlap with the teacher patterns. This shows that the dynamical behaviour of the system changes discontinuously according to the appearance of these metastable states, where

$$T_M(\alpha) \cong 1 - 1.95\alpha^{1/2} \qquad (T \sim 1, \alpha \text{ small}). \tag{4.11}$$

(e) Below $T_M(\alpha)$, the system has the line $T_c(\alpha)$ below which the retrieval solutions are globally stable, down to $T = 0$. This transition temperature $T_c(\alpha)$ is determined by equating the cost functions (free energies) of the NG and the retrieval states:

$$T_c(\alpha) \cong 1 - 2.6\alpha^{1/2} \qquad (T \sim 1, \alpha \text{ small}). \tag{4.12}$$

(f) The retrieval states become unstable to the symmetric solutions below the temperature $T_{\rm AT}(\alpha)$ which is determined from $uv = 1$ in (4.5) and (4.6) as

$$T_{\rm AT}(\alpha) \cong (8\alpha/9\pi)^{1/2} \exp(-1/2\alpha). \tag{4.13}$$

(g) The nature of the neuron system with the SB can be studied in terms of the SUSY stochastic dynamics.

(h) The non-ergodic behaviour, i.e. the chaotic behaviour appears in regimes which are unstable to the symmetric solutions. (These two results are discussed in a separate paper.)

4.1.2. General cases of a continuous neuron state. Using the expression (B.16) in place of (B.17), with the neuron state probability $h\{s\}$, e.g. with multi-peaks, we can derive similar relations to those described above but we omit them here.

4.2. Case of a synaptic junction of order $m = \lambda + 1$

Microscopic studies of neural tissue, e.g. the brain, have shown that models of neurons which interact through simple synaptic contacts with efficacy J_{ij} are oversimplified. In general, two or more axonal branches jointly contact like a dendrite and form a synaptic junction of higher order. The postsynaptic potential at time $t + 1$ is influenced if the incoming activation signals at time t are correlated. Such correlations are described in a multiplicative form of the neuron variables s_i. Then the Hopfield model is generalized to this more general situation by introducing the local field h_i:

$$h_i = \sum_{[J]} J_{ij1\ldots j\lambda} s_{j1} \ldots s_{j\lambda} \qquad (jk = j_k \text{ for } k = 1, \ldots, \lambda) \tag{4.14}$$

where $[J]$ denotes summation over $j_1 < \cdots < j_\lambda \neq i$. We study the case of the spherical model for J_{iJ}

$$\|J_i\|^2 = \sum_{[J]} J_{iJ}^2 = {}_{N-1}C_\lambda \equiv M \tag{4.15}$$

whose binomial coefficient M corresponds to the number of λ synapses coupling to a given neuron $(N - 1)$, and J_{iJ} denotes $J_{ij1\ldots j\lambda}$. The embedding condition, which guarantees the local stability of pattern μ at site i, is generalized to

$$\gamma_i^\mu[J] = (\xi_i^\mu/\|J_i\|) \sum_{[J]} J_{iJ} \xi_{j1}^\mu \ldots \xi_{j\lambda}^\mu > \kappa \qquad (jk = j_k \text{ for } k = 1, \ldots, \lambda) \tag{4.16}$$

using the threshold value κ of neurons.

We consider the limit $\beta \to \infty$ (an ideal neuron system without the influence of the environment). From the local stability condition (4.16), the cost function $H_i[J]$ and the partition function Z_i at the ith neuron are specified in (C.2) and (C.3). The OP in the replica space (C.4) is defined.

This model NNW corresponds to the spherical model [5] for $m = 2$ and the Kohring model [6] for $m = \lambda + 1$.

In order to study by means of the SUSY stochastic dynamics, a change of variables (C.5) is made. For the final expression for the partition function (C.7), using the saddle-point approximation in the limit $N \to \infty$, we found the symmetric solutions (C.8) and studied both the static and the dynamic properties of the neuron system.

Let us summarize the important results obtained.

4.2.1. Static properties.

(S1) The storage capacity (the maximum value α for a given κ) is expressed, using (C.7) and (C.12), as

$$\alpha = \lim_{N\to\infty} p/M = \left[\int_{-\kappa}^{\infty} Dt\,(t+\kappa)^2\right]^{-1}. \tag{4.17}$$

It has the famous maximum value $\alpha = 2$. This situation including (4.17) is independent of the order of interactions m.

(S2) The eigenvalues of the quadratic fluctuations around the saddle point consist of the non-degenerate eigenvalues Λ_1 and the doubly degenerate eigenvalue Λ_2 in (C.16). They are also independent of the interaction order m.

From these facts (S1) and (S2) the SUSY stochastic dynamics yield the same results as those obtained by the replica method [6], in the RS regime.

4.2.2. Dynamic properties. As the dynamics of the neuron system we specify the relation (C.17).

(D1) The retrieval OP is expressed as (C.18), which also coincides with that derived by the replica method [6].

5. Relation of information dynamics to microscopic thermal flow

Essential contributions of the information dynamics are shown to be derived from the principle of microscopic thermal flow.

For simplicity of the description we introduce the reduced cost function (free energy) W, the partition function Z and the reduced energy E for any time t, as

$$W \equiv \ln Z \qquad Z \equiv \sum \mathrm{e}^E \qquad E \equiv -\beta H = \sum_{i\neq j} J_{ij} S_i S_j + \sum_i H_i S_i \tag{5.1}$$

where J_{ij} corresponds to the Hopfield interaction ($m = 2$) in (2.2). The system potential consists of multivalleys $\{\alpha^*\}$. For each valley α^* the mean-field cost function (free energy) W and its deviation $\mathrm{d}W$ are derived as

$$\begin{aligned} W &= \sum_i A(X_i) + \sum_i (H_i - X_i) A'(X_i) + \sum_{i\neq j} J_{ij} A'(X_i) A'(X_j) \\ \mathrm{d}W &= \mu^*\,\mathrm{d}A'(X_i) + \sigma^*\,\mathrm{d}B \approx (\mu^* - \sigma^{*2}/2)\,\mathrm{d}A'(X_i) + \sigma^*\,\mathrm{d}B \end{aligned} \tag{5.2}$$

where $X_i = H_i + 2\sum_{i\neq j} J_{ij} A'(X_j)$ with the OP $A'(X_i) = M_i$ at the metastable point, $\mu^* = H_i - X_i$ and $\sigma^*\,\mathrm{d}B$ the stochastic fluctuations. Note that the valley subscript α^* is omitted for simplicity. Here the fact $\mathrm{d}A' = (\mathrm{d}B)^2$ was used based on the central limit theorem. Let us eliminate the fast stochastic fluctuations in the cost function (free energy) for the variation $\Delta A'(X_i)$ from the metastable values

$$(\partial W/\partial Z)\Delta Z - \Delta W = [-\partial W/\partial A'(X_i) - 1/2\sigma^{*2}\partial^2 W/\partial(\ln Z)^2]\Delta A'(X_i) \tag{5.3}$$

which should be proportional to $[(\partial W/\partial Z)Z - W]\Delta A'(X_i)$. Introducing the proportionality constant p^* leads to a kind of deterministic equation for the slow variables

$$\partial W/\partial A'(X_i) + 1/2\sigma^{*2}\partial^2 W/\partial(\ln Z)^2 + p^*\partial W/\partial \ln Z = p^* W. \tag{5.4}$$

By the change of variables $x \equiv A'(X_{i1}) - A'(X_i)$ and $u \equiv \ln(Z/Z_1) + (p^* - \sigma^{*2}/2)x$, the solution can be derived as

$$W = \mathrm{e}^{-p^* x} w(u, x)$$
$$w(u, x) = Z \int_{-Y2}^{\infty} Dv - Z_1 \mathrm{e}^{-p^* x} \int_{-Y1}^{\infty} Dv \tag{5.5}$$

where $\int_{-\alpha}^{\infty} Dv \equiv \int_{-\alpha}^{\infty} \exp(-v^2/2)\,\mathrm{d}v/(2\pi)^{1/2}$, $Y2 \equiv Y_2 \equiv (u + \sigma^{*2}x)/(\sigma^* x^{1/2})$, $Y1 \equiv Y_1 \equiv u/(\sigma^* x^{1/2})$ and the subscript 1 denotes the initial value. Note that the function $w(u, x)$ corresponds to the heat transfer equation

$$w_{uu} - 2\sigma^{*-2} w_x = 0. \tag{5.6}$$

Therefore, the informational flows are subject to the principle of microscopic thermal flow in each valley α^*. This fact means that certain local-temperature distributions appear in the systems. An associated multivalley potential is induced. Microscopic thermal flows are found both globally and locally. (a) The global, slow behaviour of the systems is subject to the heat transfer equation (5.6) (described with x and u above (5.5)) for the coarse-grained variables $\langle x \rangle$ and $\langle u \rangle$ over the fast variables, respectively. (b) The microscopic, slow behaviour in each valley α^* is also derived from the microscopic heat transfer equation (5.6) using the associated variables x_α and u_α coarse-grained over the scale-dependent fast variables. (c) Also for a higher step of the RBS regime, this computing process can be iterated until the neuron number contained in the associated valleys breaks the relation of the central limit theorem. That is, the method described above can be applicable to the evaluation of the systems' behaviour for any step of the RSB regime. (Of course, the systems' behaviour in the RS regime corresponds to that derived in this paper without computing iteratively.) (d) The reduced energy terms in the second relation of (5.1) are interpreted to express the energy-transfer mechanism in the heat transfer system. The information transfer of the systems is done using the principle of microscopic thermal flow. Its adaptively intelligent control can be performed by the external fields, using the mesoscopic critical behaviour for each valley α^* derived by the renormalization theory.

6. Concluding remarks

From the results obtained in the previous sections in addition to appendices A–C, the following concluding remarks are summarized.

(a) The SUSY stochastic dynamics have been established for neuron systems.
The features of the SUSY stochastic treatment are as follows.

1. Introduce both fermion and SUSY fields instead of replicating the system.
2. The stability of the system can be discussed using the FDT.
3. The contributing terms must satisfy the causality relation.
4. Within the framework of the saddle-point approximation of the model NNWs investigated in section 4, the SUSY stochastic dynamics derived the same results as those obtained by the replica method.
5. The boundary between the RS and the RSB states coincides with the $T_{\mathrm{AT}}(\alpha)$ line.
6. The SUSY stochastic treatment is very powerful for deriving solutions, they are almost as good as the rigorous solutions derived by means of the renormalization method. This is the first reason why the SUSY method is powerful.

(b) The replica treatment, inspite of the tricky treatment to replicate the system, was shown to give good results in the RS regime.

1. Both the static and dynamic behaviours are valid in the RS regime.
2. The stability of the RS states, evaluated with eigenvalues of the quadratic fluctuations around the saddle point, gives true results.
3. The boundary of the instability to the RSB is expressed as the $T_{\mathrm{AT}}(\alpha)$ line.

That is, from these facts the SUSY stochastic treatment as well as the replica treatment were shown to be very simple and powerful methods to investigate the quality of NNWs. This is the second reason for using both methods.

Below the $T_{\mathrm{AT}}(\alpha)$ line, we can expect the existence of many phases where various kinds of RSB states and low-dimensional behaviours appear. They are discussed in a separate paper by referring the characteristic scale of the associated modes as discussed in section 5. This is the third reason.

The features of the properties in the generalized NNWs of interaction order m have been investigated and summarized in section 4. Their features coincide with those obtained in the replica treatment.

A great development in the study of neurodynamics [11] was brought about by reconsidering both the slow variables $\{\xi\}$ in the heat bath T' and the fast variables $\{s\}$ in the heat bath T. Here $\{\xi\}$ denotes a set of teacher patterns and $\{s\}$ a set of neuron states. The (temperature) ratio T/T' ($\equiv \eta$) corresponds to the replica index n. The results reported in [11] coincide with those obtained by us using the SUSY method without knowledge of the existence of [11].

Finally, section 5 can be summarized as follows. (a) According to every characteristic scale of the reference modes, the information dynamics of the systems are subject to the principle of the corresponding microscopic thermal flow and fluctuate around a kind of associated deterministic behaviour. (b) This computing method is also applicable to any higher-step RSB regime and is iterated until the reference characteristic scale reaches a scale size where the central limit theorem breaks down. (c) The mesoscopic critical behaviours of the systems may be evaluated by the renormalization theory. (d) The information flows of the system can be adaptively controlled through the external fields, using the mesoscopic critical behaviours derived.

From the concluding remarks summarized above, the following points are stressed. It was shown that the SUSY and the replica formalisms yield the same, true results and that their behaviours are subject to the principle of microscopic thermal flow. The points in which the SUSY formalism is powerful and systematic are concerned with it being renormalizable and that the main correlation and response functions can be evaluated systematically according to the associated algebra. Furthermore, how to treat the RSB regime is also expected to be solvable by the similarity or conformal transformation. In this process the renormalization is also applicable.

Acknowledgments

The author would like to thank the referees for careful advice and comments including information concerning reference [11].

Appendix A. SUSY stochastic dynamics

By introducing $p_i \equiv -\mathrm{i}\partial/\partial s_i$ in (3.3), the time evolution of the probability $P(s,t)$ ($s \equiv \{s\}$) is expressed with the Fokker–Planck equation [3]

$$\partial P/\partial t = -\Gamma_0 \sum_i p_i[p_i - \mathrm{i}\beta\partial E/\partial s_i] \equiv -H_{\mathrm{FP}}P \tag{A.1}$$

where the non-Hermitian FP operator is written as

$$H_{\mathrm{FP}} = \Gamma_0 \sum_i B_i^+ B_i^- \tag{A.2}$$

by formally introducing the creation and annihilation operators $B_i^+ \equiv p_i$ and $B_i^- \equiv p_i - \mathrm{i}\beta\partial E/\partial s_i$. Here note that these two operators are not Hermitian to each other. Using a (non-unitary) transformation matrix T^* ($\equiv \exp\{-\beta E(s)/2\}$), the following Hermitian forms are obtained:

$$\begin{aligned} B_i^{*\pm} &\equiv T^{*-1} B_i^{\pm} T^* = p_i \pm (\mathrm{i}\beta/2)\partial E/\partial s_i \\ H_{\mathrm{FP}}^* &\equiv T^{*-1} H_{\mathrm{FP}} T^* = \Gamma_0 \sum_i B_i^{*+} B_i^{*-}. \end{aligned} \tag{A.3}$$

As a result, the eigenvalues of H_{FP}^* are non-negative. For zero eigenvalues, the right (left) eigenvector of H_{FP} is proportional to the canonical weight $P_{eq} = \mathrm{e}^{-\beta E}$ (1, respectively).

In order to be able to choose randomly one pattern out of the pattern set $\{\xi\}$, introduce $2N$ fermions $\{a_i^+, a_i\}$ and SUSY operators

$$Q^+ \equiv \sum_i B_i^- a_i^+ \qquad Q^- \equiv \sum_i B_i^+ a_i \tag{A.4}$$

and define

$$Q_1 \equiv Q^+ + Q^- \qquad Q_2 \equiv (Q^+ - Q^-)/\mathrm{i} \qquad N_F \equiv \sum_i a_i^+ a_i. \tag{A.5}$$

Then the following relations are derived:

$$\begin{aligned} &(Q^\pm)^2 \equiv 0 \qquad [N_F, Q^\pm]_- = \pm Q^\pm \\ &H_{\mathrm{FP}} = \Gamma_0\left[\sum_i p_i(p_i - \mathrm{i}\beta\partial E/\partial s_i) + \sum_{ij} a_i^+ a_j \beta \partial^2 E/\partial s_i \partial s_j\right]. \end{aligned} \tag{A.6}$$

Within the framework of the SUSY fields, it is easily shown that the following (charge and fermion number) conservation relations hold:

$$[H_{\mathrm{FP}}, Q^\pm]_- = 0 \qquad [H_{\mathrm{FP}}, N_F]_- = 0 \tag{A.7}$$

and that the actual eigenvectors are described as

$$\begin{aligned} |0\rangle &\equiv N^{-1}\mathrm{e}^{-\beta E(s)} \otimes |F0\rangle \\ \langle 0| &\equiv N^{-1} 1 \otimes (a-b)\langle F0| \end{aligned} \tag{A.8}$$

where N^{-1} denotes the normalization constant, $\otimes$ the direct product, and $|F0\rangle$ ($\langle F0|$) the right (left) eigenvector for the fermion fields with zero eigenvalue (respectively). Note that the expectation values of $\{Q^\pm, N_F\}$ are always conserved.

According to quantum theory, correlation functions for neurons are expressed as

$$\langle O_a(t+\tau)O_b(t)\rangle_{\mathrm{T}} = \mathrm{tr}\left[\mathrm{e}^{\mathrm{i}\pi NF}\mathrm{e}^{-TH}O_a\mathrm{e}^{-\tau H}O_b\mathrm{e}^{\tau H}\right] = \sum_{ij}(-1)^{NFi}\mathrm{e}^{-T\varepsilon i-\tau(\varepsilon j-\varepsilon i)}\langle i|O_a|j\rangle\langle j|O_b|i\rangle \tag{A.9}$$

where $\langle\ \rangle_{\mathrm{T}}$ denotes the environmental (thermal) average, H abbreviates H_{FP} and $\varepsilon k = \varepsilon_k$ for $k = i, j$.

The time evolution of these correlation functions is also derived as follows:

$$\partial/\partial\tau\langle O_a(t+\tau)O_b(t)\rangle_{\mathrm{T}} = \Gamma_0\,\mathrm{tr}[\mathrm{e}^{\mathrm{i}\pi NF}\mathrm{e}^{-TH}O_a\mathrm{e}^{-\tau H}[[O_b, Q^+]_-, Q^-]_+\mathrm{e}^{\tau H}] - \Gamma_0\,\mathrm{tr}[\mathrm{e}^{\mathrm{i}\pi NF}\mathrm{e}^{-TH}[[O_a, Q^+]_-, Q^-]_+\mathrm{e}^{-\tau H}O_b\mathrm{e}^{\tau H}] \qquad (NF = N_F) \tag{A.10}$$

in particular, for $O_a = O_b \equiv s_k$

$$\partial/\partial\tau\langle s_k(t+\tau)\,s_k(t)\rangle_{\mathrm{T}} = \mathrm{i}\Gamma_0[\langle s_k(t+\tau)\,p_k(t)\rangle_{\mathrm{T}} - \langle p_k(t+\tau)\,s_k(t)\rangle_{\mathrm{T}}].$$

These relations are called the fluctuation–dissipation theorem (FDT).

From the properties of the ground state ($\langle 0|p_i = \langle 0|a_i^+ = 0$), it is also shown that the following causality relations hold for any operation O:

$$\begin{aligned}&\langle p_i(t+\tau)\,O(t)\rangle = \langle a_k^+(t+\tau)\,O(t)\rangle = 0\\&\lim_{\tau\to 0+}\langle s_i(t+\tau)\,p_i(t)\rangle = \mathrm{i}\\&\lim_{\tau\to 0+}\langle a_i(t+\tau)\,a_i^+(t)\rangle = 1.\end{aligned} \tag{A.11}$$

By applying the method of path integration to the FP equation in the superspace, the partition functions are obtained:

$$\begin{aligned}Z &= \int \mathrm{e}^{-s}D[s]\,D[p]\,D[\eta]\,D[\bar\eta]\\ S &= \int_0^T \mathrm{d}t\left[\Gamma_0^{-1}\sum_i([\dot\eta_i\bar\eta_i + p_i\dot s_i - p_i^2) + \sum_i(\partial E/\partial s_i)p_i + \sum_{i,j}(\partial^2E/\partial s_i\partial s_j)\eta_i\bar\eta_j\right]\end{aligned} \tag{A.12}$$

where $\eta, \bar\eta$ are fermionic variables ($\eta\eta = \bar\eta\bar\eta = 0$, $\eta_i\bar\eta_j = -\bar\eta_j\eta_i$ $(i \neq j)$, etc) and the dot denotes a time derivative.

Let us define the superfield ϕ_i and its arguments a in the superspace as

$$\phi_i \equiv s_i + \bar\theta\eta_i + \bar\eta_i\theta + p_i\bar\theta\theta \qquad a \equiv (t_a, \dot{\bar\theta}_a, \theta_a) \tag{A.13}$$

using anticommuting Grassmann variables $\bar\theta, \theta$. Here the superspace delta and the Grassmann delta functions are defined, respectively, as

$$\delta(a-b) = \delta(t_a - t_b)\delta^2(\Theta_a - \Theta_b) \qquad \delta^2(\Theta_a - \Theta_b) = (\bar\theta_a - \bar\theta_b)(\theta_a - \theta_b). \tag{A.14}$$

Then the final expression for the SUSY stochastic dynamics is expressed as follows:

$$\begin{aligned}Z &= \int D[\phi]\mathrm{e}^{-Sk-Sp}\\ Sk \equiv S_k &\equiv -\tfrac{1}{2}\int \mathrm{d}\theta\,\mathrm{d}\bar\theta\,\mathrm{d}t\sum_i \phi_i D^{(2)}\phi_i\\ Sp \equiv S_p &\equiv \int \mathrm{d}\theta\,\mathrm{d}\bar\theta\,\mathrm{d}t\,E(\phi)\end{aligned} \tag{A.15}$$

with

$$D^{(2)} \equiv 2\partial^2/\partial\theta\partial\bar{\theta} + 2\theta\partial^2/\partial\bar{\theta}\partial t - \partial/\partial t. \tag{A.16}$$

All neuron correlation and response functions of the system are included in the SUSY correlation function

$$\langle Q(a,b)\rangle = 1/N \sum_i \langle \phi_i(a)\, \phi_i(b)\rangle \tag{A.17}$$

because of the definition of ϕ_i in (A.13).

Appendix B. Derivation in the case of a synaptic junction of order $m = 2$

We consider the neural system with $m = 2$ in the expression (2.1). To linearize the partition function of the system with respect to ξ for the sake of the trace on $\{\xi\}$, the Gaussian integrals are used:

$$\int D[m] \exp\left[-\beta N \left\{ \tfrac{1}{2}(m^\mu)^2 - m^\mu 1/N \sum_i \xi_i^\mu s_i \right\}\right] = \exp\left[(\beta/N)\sum_{i<j} \xi_i^\mu \xi_j^\mu s_i s_j\right]. \tag{B.1}$$

In this step the neuron states $\{s\}$ are replaced by the SUSY fields $\{\phi\}$ in (A.13), as

$$\begin{aligned}(\text{B.1}) \to \int D[M] \exp\bigg[-\beta N \bigg\{ \tfrac{1}{2} \int \mathrm{d}a\, \mathrm{d}b\, M^\mu(a) M^\mu(b) \delta(a-b) \\ - \int \mathrm{d}a\, M^\mu(a) 1/N \sum_i \xi_i^\mu \phi(a) \bigg\}\bigg]\end{aligned} \tag{B.2}$$

and the constraints for the OPs $\{r_{ab}, q_{ab}\}$ and for the spherical neuron states are imposed:

$$\begin{aligned}&\int D[R^*/(2\pi)] D[R] \exp\left[-\mathrm{i}/2 \int \mathrm{d}a\, \mathrm{d}b\, R^*(a,b) \left\{ R(a,b) - N/p \sum_{\mu>1} M^\mu(a) M^\mu(b) \right\}\right] \\ &\int D[Q^*/(2\pi)] D[Q] \exp\left[-\mathrm{i}/2 \int \mathrm{d}a\, \mathrm{d}b\, Q^*(a,b) \left\{ Q(a,b) - 1/N \sum_i \phi_i(a)\phi_i(b) \right\}\right] \\ &\int D[\bar{\bar{Q}}] \exp\left[-\mathrm{i}/2 \int \mathrm{d}a\, Q(a) \{\bar{\bar{Q}}(a,a) - 1\}\right].\end{aligned} \tag{B.3}$$

Then, by taking into account averages over both the recalling pattern ($\mu = 1$) and the non-recalling ones ($\mu = 2, 3, \ldots, p$), the leading expression for the partition function in N is obtained as follows:

$$\begin{aligned} Z &= \left\langle\!\!\left\langle \int D[M^1] D[R^*/(2\pi)] D[R] D[Q^*/(2\pi)] D[Q] D[Q] \mathrm{e}^{-s} \right\rangle\!\!\right\rangle_{\{\xi\}} \\ S/(\beta N) &= \tfrac{1}{2} \int \mathrm{d}a\, \mathrm{d}b\, M^1(a) M^1(b) \delta(a-b) + (p-1)/(2\beta N)\, \mathrm{tr} \ln \Lambda^* \\ &\quad + \mathrm{i}/(2\beta N) \int \mathrm{d}a\, \mathrm{d}b\, R^*(a,b) R(a,b) + \mathrm{i}/(2\beta N) \int \mathrm{d}a\, \mathrm{d}b\, Q^*(a,b) Q(a,b) \\ &\quad - \mathrm{i}/(2\beta N) \int \mathrm{d}a\, \bar{\bar{Q}}(a) [1 - Q(a,a)] - W/(\beta N) \end{aligned} \tag{B.4}$$

where

$$\exp(-W) \equiv \left\langle\!\!\left\langle \int D[\phi] \exp\left[\beta N \left\{ \int \mathrm{d}a\, M^1(a)\phi(a)\xi^1 - 1/(2\beta N) \int \mathrm{d}a\,\mathrm{d}b\, \phi(a)\kappa(a,b)\phi(b) \right\}\right]\right\rangle\!\!\right\rangle_{\xi 1} \tag{B.5}$$

$$\kappa(a,b) \equiv \Gamma_0^{-1} D^{(2)}(a)\delta(a-b) - \mathrm{i}\bar{\bar{Q}}(a,b)$$

$$\bar{\bar{Q}}(a,b) \equiv Q^*(a,b) - \mathrm{i}\beta^2 N \sum_{\mu>1} M^\mu(a) M^\mu(b)$$

and

$$\int D[M^\mu] \exp\left[(-\beta N/2) \int \mathrm{d}a\,\mathrm{d}b\, M^\mu(a)\{(1-\beta)\delta(a-b) - \beta Q(a,b) - \mathrm{i}/(\beta p) R^*(a,b)\} M^\mu(b)\right]$$
$$\equiv \int D[M^\mu] \exp[(-\beta N/2) \int \mathrm{d}a\,\mathrm{d}b\, M^\mu(a)\Lambda(a,b)M^\mu(b)] \qquad (\mu > 1) \tag{B.6}$$

i.e.

$$\Lambda(a,b) \equiv (1-\beta)\delta(a-b) - \beta Q(a,b) - \mathrm{i}/(\beta p) R^*(a,b). \tag{B.7}$$

Using the relation (B.4), the neuron correlation functions are evaluated as

$$\langle Q(a,b)\rangle = \int D[Q] Q(a,b) \mathrm{e}^{-s} = \mathrm{i}2\partial \ln Z/\partial Q^*(a,b).$$

We consider the evaluation of S in the large-N limit, i.e. the saddle-point approximation. If there is a single saddle point of the system the correlation functions take the saddle-point value

$$\lim_{N\to\infty} \langle Q(a,b)\rangle = \langle Q(a,b)\rangle|_{\text{saddle point}}. \tag{B.8}$$

At the saddle point the system has the following extremum values:

(a) $\partial S/\partial M^1(a) = 0 \qquad M^1(a) = \langle\!\langle \xi^1 \langle\phi(a)\rangle\rangle\!\rangle$

(b) $\partial S/\partial \bar{Q}(a,b) = 0 \qquad Q(a,b) = \langle\!\langle \langle\phi(a)\rangle\langle\phi(b)\rangle \rangle\!\rangle \qquad (a \neq b)$

(c) $\partial S/\partial Q(a,b) = 0 \qquad \bar{Q}(a,b) = \mathrm{i}(p-1)/2\langle\!\langle \partial\, \mathrm{tr} \ln \Lambda^*/\partial Q(a,b)\rangle\!\rangle$

$$= -\mathrm{i}\beta^2 N \left\langle\!\!\left\langle \sum_{\mu>1} M^\mu(a)M^\mu(b) \right\rangle\!\!\right\rangle = -\mathrm{i}p\beta^2 R(a,b) \qquad (a \neq b) \tag{B.9}$$

(d) $\partial S/\partial R^*(a,b) = 0 \qquad R(a,b) = N/p \left\langle\!\!\left\langle \sum_{\mu>1} M^\mu(a)M^\mu(b) \right\rangle\!\!\right\rangle \qquad (a \neq b)$

(e) $\partial S/\partial R(a,b) = 0 \qquad R^*(a,b) = 0 \qquad (a \neq b)$

(f) $\partial S/\partial Q(a,b) = 0 \qquad Q(a,a) = 1$

where the average over the recalling pattern ξ^1 is denoted by

$$\langle\!\langle\ \rangle\!\rangle = \langle\!\langle\ \rangle\!\rangle_{\xi 1}. \tag{B.10}$$

The first relation (a) expresses the retrieval OP, i.e. the overlap between the recalling pattern and the mean neuron states; (b) denotes the neuron glass OP, i.e. the overlap of the averaged neuron states; (d) denotes the overlap of the retrievals. The last relation in (c) is derived using expression (d). Using relation (e) the transfer matrix (B.7) is rewritten as

$$\Lambda_0(a,b) \equiv (1-\beta)\delta(a-b) - \beta Q(a,b). \tag{B.11}$$

Relation (f) expresses the spherical neuron states.

Now let us mention a causal-supersymmetric delta function $\delta_{cs}(a-b)$ written in (B.11). This function is defined for any function $F(|\tau|)$ such as $F(0)=1$ and $\lim_{\tau\to\infty} F(|\tau|)=0$:

$$\delta_{cs}(a-b) = \lim_{\Delta\to\infty}\left[1+\tfrac{1}{2}(\bar{\theta}_a-\bar{\theta}_b)\{\theta_a+\theta_b-(\theta_a-\theta_b)\Xi(t_a-t_b)\}\partial/\partial t_a\right]F(\Delta|t_a-t_b|) \tag{B.12}$$

using the relation $a \equiv (t_a, \bar{\theta}_a, \theta_a)$ and the sign function of τ, $\Xi(\tau)$. For example, $\delta_{cs}=1$ for $a=b$.

We define

$$V_o(a,b) \equiv \delta_{cs}(a-b) \qquad V_1(a,b) \equiv 1 \tag{B.13}$$

and consider the symmetric solutions

$$m \equiv M^1(a) \qquad q \equiv Q(a,b) \qquad r \equiv R(a,b) \tag{B.14}$$

independent of the variables a, b.

The matrix (B.11) and its inverse matrix are expressed as follows:

$$\begin{aligned}\Lambda_0(a,b) &= (1-\beta)V_0(a,b) - \beta q[V_1(a,b)-V_0(a,b)]\\ \Lambda_0^{-1}(a,b) &= [1-\beta(1-q)]^{-2}[\{1-\beta(1-2q)\}V_0(a,b)+\beta q\{V_1(a,b)-V_0(a,b)\}].\end{aligned} \tag{B.15}$$

The expression (B.5) is written as

$$\begin{aligned}\exp(-W_0) \equiv \left\langle\!\!\left\langle \int D[\phi]\exp\left[\beta N\left\{\int \mathrm{d}a\, M^1(a)\phi(a)\xi^1\right.\right.\right.\right.\\ \left.\left.\left.\left. +(\alpha\beta r/2)\int \mathrm{d}a\,\mathrm{d}b\,\phi(a)\phi(b) - (\alpha\beta r/2)\int \mathrm{d}a\,\phi^2(a)\right\}\right]\right\rangle\!\!\right\rangle.\end{aligned} \tag{B.16}$$

As a simple case, consider the discrete limit of neuron states $S=\pm 1$. The expression (B.16) is rewritten as

$$\begin{aligned}\exp(-W_0) &\equiv \left[\left\langle\!\!\left\langle \sum_{[\phi]}\exp\left\{\beta\int \mathrm{d}a\, M^1(a)\phi(a)\xi^1\right.\right.\right.\right.\\ &\qquad \left.\left.\left.\left. +(\alpha\beta^2 r/2)\int \mathrm{d}a\,\mathrm{d}b\,\phi(a)\phi(b))\right\}\right\rangle\!\!\right\rangle \exp(-\alpha\beta^2 r/2)\right]^N\\ &= \left[\left\langle\!\!\left\langle \int D[z]\exp\{-1/2z^2 + \mathrm{tr}\ln\mathrm{ch}\,\beta[z(\alpha r)^{1/2}+m^1\xi^1]\}\right\rangle\!\!\right\rangle \exp(-\alpha\beta^2 r/2)\right]^N.\end{aligned} \tag{B.17}$$

The extremum conditions of S for the set of OPs $\{m^1, r, q\}$ yield the set of coupled equations for OPs $\{m^1, r, q\}$, i.e.

(a′) $\qquad \partial S/\partial m^1 = 0 \qquad m = \int Dz\langle\!\langle \xi\ \mathrm{th}[\beta(z(\alpha r)^{1/2}+\xi m)]\rangle\!\rangle$

$$\text{(b}'\text{)} \qquad \partial S/\partial r = 0 \qquad q = 1 - (T/\alpha)\int Dzz\langle\!\langle(\alpha/r)^{1/2}\,\mathrm{th}[\beta(z(\alpha r)^{1/2} + \xi m)]\rangle\!\rangle$$

$$= \int Dz\langle\!\langle \mathrm{th}^2\,\beta[(z(\alpha r)^{1/2} + \xi m)]\rangle\!\rangle$$

$$\text{(c}'\text{)} \qquad \partial S/\partial q = 0 \qquad r = q/[1 - \beta(1-q)]^2 \tag{B.18}$$

where $\{m^1, \xi^1\}$ are denoted by $\{m, \xi\}$, and $Dz \equiv \exp(-z^2/2)\,\mathrm{d}z/(2\pi)^{1/2}$. The first relation (a′) is easily derived from differentiation of (B.4) and (B.17). The second one (b′) is also obtained from the differentiations of the fourth term on the right-hand side of (B.4) and the terms on right-hand side of (B.17). The last relation (c′) is derived by inserting the term Λ_0^{-1} and the derivative $\partial\Lambda_0/\partial q$ of (B.15) in expression (c) of (B.9).

The following average calculation over the recalling pattern ξ is easily performed by making a change of variables $z \to \xi z$. The final expressions are as follows:

$$\begin{aligned} m &= \int Dz\,\mathrm{th}\,\beta(z(\alpha r)^{1/2} + m) \\ q &= \int Dz\,\mathrm{th}^2\,\beta(z(\alpha r)^{1/2} + m). \end{aligned} \tag{B.19}$$

Finally, we derive the eigenvalues of the quadratic fluctuations around the saddle point. The differentiation of S in (B.4) with respect to $Q(a, b)$ leads to the relation

$$\Gamma_0^{-1} D^{(2)}(a)\delta(a-b) + \mathrm{i}\bar{\bar{Q}}(a)\delta(a-b) + Q^{-1}(a,b) = 0 \tag{B.20}$$

where we use the fact that the term W is expressed as a sum over connected diagrams with propagator κ^{-1} and that the stationarity of S with respect to $Q(a, b)$ gives relation (b) of (B.10);

$$Q(a,b) = \mathrm{i}\delta W/\delta\bar{Q}(a,b) = \langle\phi(a)\,\phi(b)\rangle. \tag{B.21}$$

Neglecting the kinetic term, we multiply both sides of (B.20) by $\delta(a - b)$:

$$-\mathrm{i}\delta(a-b)\bar{\bar{Q}}(a) = \delta(a-b)Q^{-1}(a,b).$$

Substituting this expression into (B.20) and differentiating with respect to $Q(c, d)$ leads to the eigenvalue equation

$$\begin{aligned} &(\Lambda_1 - \Lambda)\delta q(a-b) + c_1c_2\left[\int \mathrm{d}c\,\delta q(a,c) + \int \mathrm{d}d\,\delta q(d,b)\right] + c_2^2\int \mathrm{d}c\,\mathrm{d}d\,\delta q(c,d) \\ &\qquad -\delta(a-b)\left[2c_1c_2\int \mathrm{d}c\,\delta q(a,c) + c_2^2\int \mathrm{d}c\,\mathrm{d}d\,\delta q(c,d)\right] = 0 \end{aligned} \tag{B.22}$$

where

$$\begin{aligned} &c_1 \equiv \beta[1-\beta(1-q)]^{-1} \qquad && c_2 \equiv \beta q[1-\beta(1-q)]^{-2} \\ &\delta q(a,a) = 0 && \Lambda_1 \equiv \beta[1-\beta(1-q)]^{-2} \end{aligned} \tag{B.23}$$

using (B.15). The relation (B.22) gives the following three eigenvalues:

$$\begin{aligned} &\Lambda_1 && \text{for (a)} \quad \int \mathrm{d}a\,\delta q(a,b) = \int \mathrm{d}a\,\delta q(b,a) = 0 \\ &\Lambda_2 \equiv \Lambda_1 - 2c_1c_2 && \text{for (b)} \quad \int \mathrm{d}a\,\mathrm{d}b\,\delta q(a,b) = 0 \quad \text{except (a)} \\ &\Lambda_3 \equiv \Lambda_2 && \text{for (c)} \quad \int \mathrm{d}a\,\mathrm{d}b\,\delta q(a,b) \neq 0 \quad \text{except (a).} \end{aligned} \tag{B.24}$$

In general, the stability matrix for the symmetric solutions has the general structure

$$M = \begin{pmatrix} A(ab, cd) & C(ab, cd) \\ C(ab, cd) & B(ab, cd) \end{pmatrix}$$

where

$$\begin{aligned} A(ab, cd) &= \partial^2 S/\partial Q(a, b)\partial Q(c, d) \\ B(ab, cd) &= \partial^2 S/\partial R(a, b)\partial R(c, d) \\ C(ab, cd) &= \partial^2 S/\partial Q(a, b)\partial R(c, d) \end{aligned} \tag{B.25}$$

which are completely similar to those obtained in the replica method [4]. In a similar way, the stability of the system can be discussed using the relation

$$\delta Q(a, b) = D(a, b) \qquad \delta R(a, b) = \kappa(a, b) \tag{B.26}$$

under the constraint

$$\int \mathrm{d}b\, D(a, b) = 0 \qquad \text{for all } a. \tag{B.27}$$

Finally, the eigenvalues λ are obtained as

$$\begin{aligned} \lambda &= 2\alpha\beta\tilde{\lambda} \\ \tilde{\lambda}_\pm &= -1/2(u + v) \pm [1/4(u + v)^2 + 1 - uv]^{1/2} \end{aligned} \tag{B.28}$$

where

$$u \equiv \alpha\beta^2 \langle\!\langle (1 - \langle s\rangle^2)^2 \rangle\!\rangle \qquad v \equiv [1 - \beta(1 - q)]^{-2}. \tag{B.29}$$

When the inverse environment parameter, i.e. temperature T is higher than the glass-phase temperature T_g $(= 1 + \alpha^{1/2})$, the OP q is null and $uv = \alpha\beta^2/(1 - \beta)^2 < 1$. Therefore, λ_- is negative and λ_+ positive in this regime. The change of sign of λ_+ corresponds to the RSB, which occurs on the boundary of $uv = 1$. These results coincide with those obtained in the replica method.

Appendix C. Derivation in the case of a synaptic junction of order $m = \lambda + 1$

In this neuron model the postsynaptic state at time $t + 1$ is supposed to be determined by the local field at time t in (4.14):

$$S_i(t + 1) = \mathrm{sgn}(h_i(t)) \tag{C.1}$$

for the ith neuron. Using the local stability condition (4.16) the cost function $H_i[J]$ and the partition function Z_i at the ith neuron are expressed as follows:

$$\begin{aligned} H_i[J] &= \sum_{\mu=1}^{p} \theta(\kappa - \gamma_i^\mu[J]) \\ Z_i &= \left\langle\!\!\left\langle \int D[J]\rho[J] \prod_{\mu=1}^{p} \theta(\kappa - \gamma_i^\mu[J]) \right\rangle\!\!\right\rangle_{\{\xi\}}. \end{aligned} \tag{C.2}$$

The density of state $\rho[J]$ is given by

$$\rho[J] = \delta\Big(M^{-1} \sum_{[J]} J_{iJ}^2 - 1\Big) \Big/ \int D[J]\delta\Big(M^{-1} \sum_{[J]} J_{iJ}^2 - 1\Big) \tag{C.3}$$

and averaging over the patterns $\{\xi\}$, $\langle\!\langle\ \rangle\!\rangle_{\{\xi\}}$, leads to the OP

$$q^{\alpha\beta} = M^{-1} \sum_{[J]} J^{\alpha}_{iJ} J^{\beta}_{iJ} \tag{C.4}$$

in the replica method.

In order to transform these relations into the superspace representation, the following replacements:

$$\begin{aligned} &J_{iJ} \to \phi_J \\ &q^{\alpha\beta} \to Q(a,b) = M^{-1} \sum_{[J]} \phi_J(a)\phi_J(b) \qquad N \to M \end{aligned} \tag{C.5}$$

and the constraint for $Q(a,b)$ are used. The partition function is rewritten as

$$\begin{aligned} Z_i = \Bigg\langle\!\!\Bigg\langle \int & D[\phi] D[\bar{\bar{Q}}/(2\pi)] D[Q] D[Q^*/(2\pi)] D[x] D[x^*/(2\pi)] \\ & \times \exp\Bigg[\mathrm{i} \int \mathrm{d}a \sum_{\mu} x^{*\mu}(a) \Bigg\{ x^{\mu}(a) - M^{-1/2} \sum_{[J]} \phi_J(a) \xi^{\mu}_{j1} \ldots \xi^{\mu}_{j\lambda} \Bigg\} \\ & + \mathrm{i} \int \mathrm{d}a\, Q(a) \Bigg\{ M^{-1} \sum_{[J]} \phi_J^2(a) - 1 \Bigg\} \\ & + \mathrm{i}/2 \int \mathrm{d}a\, \mathrm{d}b\, Q^*(a,b) \Bigg\{ Q(a,b) - M^{-1} \sum_{[J]} \phi_J(a)\phi_J(b) \Bigg\} \Bigg] \Bigg\rangle\!\!\Bigg\rangle_{\{\xi\}} \end{aligned} \tag{C.6}$$

where $jk \equiv j_k$ for $k = 1, \ldots, \lambda$. Using the identity $\exp(-\mathrm{i}\xi x) = \cos x(1 - \mathrm{i}\xi \tan x)$, taking the trace over ξ and approximating $\ln \cos x \approx \ln(1 - x^2/2) \approx -x^2/2$, we can express the leading terms of Z_i in N, as

$$\begin{aligned} & Z_i = \int D[\bar{\bar{Q}}/(2\pi)] D[Q] D[Q^*/(2\pi)] \exp(-S) \\ & S \equiv S_1 - \mathrm{i}/2 \int \mathrm{d}a\, \mathrm{d}b\, Q^*(a,b) Q(a,b) + \mathrm{i} \int \mathrm{d}a\, \bar{\bar{Q}}(a) + W \\ & S_1 \equiv -p \ln \int_{\kappa}^{\infty} D[x/(2\pi)] D[x^*] \exp\Big[- \int \mathrm{d}a\, \mathrm{d}b\, \big\{ \tfrac{1}{2} x^*(a) x^*(b) \delta(a-b) \\ & \qquad + \tfrac{1}{2} Q(a,b) x^*(a) x^*(b) - \mathrm{i} x^*(a) x(a) \delta(a-b) \big\} \Big] \\ & \mathrm{e}^W \equiv \int D[\phi] \exp[-1/(2M) \sum_{[J]} \int \mathrm{d}a\, \mathrm{d}b\, \phi_J(a) \kappa^*(a,b) \phi_J(b)] \\ & \kappa^*(a,b) \equiv \Gamma_0^{-1} D^{(2)}(a) = \delta(a-b) + \mathrm{i} Q^*(a,b) - \mathrm{i}2 Q(a) \delta(a-b). \end{aligned} \tag{C.7}$$

In the limit $N \to \infty$, using the saddle-point method, consider the symmetric solutions

$$\bar{\bar{Q}} \equiv \bar{\bar{Q}}(a) \qquad Q^* \equiv Q^*(a,b) \qquad q \equiv Q(a,b). \tag{C.8}$$

Neglecting the first term in $\kappa^*(a,b)$, integrate over ϕ in W:

$$\begin{aligned} & W \approx \ln(\det \Lambda_1)^{-1/2} \\ & \Lambda_1 \equiv -\mathrm{i}2 \bar{\bar{Q}} V_0(a,b) + \mathrm{i} Q^* [V_1(a,b) - V_0(a,b)] \\ & \Lambda_1^{-1} \equiv (2\bar{\bar{Q}} + Q^*)^{-2} [\mathrm{i}(2\bar{\bar{Q}} + 2Q^*) V_0(a,b) + \mathrm{i} Q^* \{V_1(a,b) - V_0(a,b)\}]. \end{aligned} \tag{C.9}$$

Next, integrate over x^* in S_1:

$$-S_1/p = \ln \int_\kappa^\infty D[x/(2\pi)](\det \Lambda_2)^{-1/2} \exp\left[-\tfrac{1}{2}\int \mathrm{d}a\,\mathrm{d}b\, x(a)\Lambda_2^{-1}(a,b)x(b)\right]$$

where

$$\begin{aligned}\Lambda_2(a,b) &= V_0(a,b) + q[V_1(a,b) - V_0(a,b)]\\ \Lambda_2^{-1}(a,b) &= (1-q)^{-2}[(1-2q)V_0(a,b) - q\{V_1(a,b) - V_0(a,b)\}].\end{aligned} \tag{C.10}$$

Furthermore, the exponential term is evaluated as

$$\begin{aligned}&\int_\kappa^\infty D[x/(2\pi)] \exp\bigg[-\tfrac{1}{2}\int \mathrm{d}a\,\mathrm{d}b\, x(a)x(b)\delta_{cs}(a-b)(1-q)^{-1}\\ &\qquad +\tfrac{1}{2}\left(\int \mathrm{d}a\, x(a)\right)^2 q(1-q)^{-2}\bigg]\\ &\quad = \int_\kappa^\infty D[x/(2\pi)] \int D[z/(2\pi)] \exp\bigg[-\tfrac{1}{2}\int \mathrm{d}a\,\mathrm{d}b\, x^2(a)\delta_{cs}(a-b)(1-q)^{-1}\\ &\qquad -\int \mathrm{d}a\, x(a) z q^{1/2}(1-q)^{-1} - \tfrac{1}{2}z^2\bigg]\\ &\quad \approx \int_{-\infty}^\infty \mathrm{d}z/(2\pi)^{1/2} \exp[-z^2/2] \int_\kappa^\infty D[x/(2\pi)]\\ &\qquad \times \exp\left[-\{2(1-q)\}^{-1}\left\{\int \mathrm{d}a\, x(a) + q^{1/2}z\right\}^2\right].\end{aligned} \tag{C.11}$$

Therefore, the following relation is obtained:

$$-S_1/p = \ln\left[(\det \Lambda_2)^{-1/2}\int_{-\infty}^\infty Dz\, H(\tau)\right] \tag{C.12}$$

using

$$t \equiv [q^{1/2}z + x]/(1-q)^{1/2} \qquad Dz \equiv \exp(-z^2/2)\,\mathrm{d}z/(2\pi)^{1/2}$$

$$\tau \equiv (q^{1/2}z + \kappa)/(1-q)^{1/2} \qquad H(\tau) \equiv \int_\tau^\infty Dt$$

where $H(x) \equiv \frac{1}{2}\operatorname{erfc}(x/\sqrt{2})$ and the complementary error function is related to the error function as $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$.

The saddle point is given by the following extremum conditions:

$$\partial S/\partial Q^*(a,b) = \partial S/\partial \bar{\bar{Q}}(a) = \partial S/\partial Q(a,b) = 0. \tag{C.13}$$

The first and the second conditions give a set of coupled equations

$$\mathrm{i} + A + FB = 0 \qquad \mathrm{i}q + A - 2EB = 0 \tag{C.14}$$

using $A \equiv (2\bar{\bar{Q}} + Q^*)^{-1}$ and $B \equiv A^2$. Their solutions are obtained as follows:

$$2\bar{\bar{Q}} = \mathrm{i}(1-2q)(1-q)^{-2} \qquad Q^* = \mathrm{i}q(1-q)^{-2}. \tag{C.15}$$

The third condition in (C.13) leads, in the same way as (B.20)–(B.24), to

$$\Gamma_0^{-1}D^{(2)}(a)\delta(a-b) - \mathrm{i}2\bar{\bar{Q}}(a)\delta(a-b) + Q^{-1}(a,b) = 0$$

which has the following eigenvalues of the quadratic fluctuations around the saddle point:

$$\begin{aligned} \Lambda_1 &= (1-q)^{-2} && \text{for (a) in (B.24)} \\ \Lambda_2 = \Lambda_3 &= \Lambda_1 - 2q(1-q)^{-3} && \text{for (b) and (c) in (B.24).} \end{aligned} \tag{C.16}$$

In order to understand the dynamic behaviour of the system ($\lambda > 1$), we consider the same probability $P(m(a)|m(b))$ as that in the Kohring model [6]. It is the probability of the system transferring the state with overlap $m(a)$ at time $t = a$ from $m(b)$ at time $t = b$. Here we suppose that $\{m(a), m(b)\}$ denote the overlap of a neuron state with one of the teacher patterns, i.e. the recalling pattern and that the other patterns are the non-recalling ones:

$$\begin{aligned} P(m(a)|m(b)) = \int D[s_i(b)]\delta\left(m(a) - N^{-1}\sum_i \mathrm{sgn}\{\xi_i h_i(b)\}\right) \\ \times\delta\left(m(b) - N^{-1}\sum_j \xi_j s_j(b)\right) \Big/ \int D[s_i(b)]\delta\left(m(b) - N^{-1}\sum_j \xi_j s_j(b)\right). \end{aligned} \tag{C.17}$$

Using the SUSY stochastic dynamics we obtain the final result

$$\begin{aligned} m(a) = \int_\kappa^\infty Dz\, \mathrm{erf}(zm^\lambda(b)/\{2(1-m^{2\lambda}(b))\}^{1/2}) \\ +\tfrac{1}{2}\,\mathrm{erf}(\kappa m^\lambda(b)/\{2(1-m^{2\lambda}(b))\}^{1/2})\,\mathrm{erfc}(-\kappa/2^{1/2}) \end{aligned} \tag{C.18}$$

whose derivation is omitted here because the method used is similar to that mentioned in appendix B and the calculation process is too long. Expression (C.18) coincides with that derived by means of the replica method [6].

References

[1] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)

[2] Müller B and Reinhardt J 1990 *Neural Networks—an Introduction* (Berlin: Springer)

[3] Zinn-Justin J 1989 *Quantum Field Theory and Critical Phenomena* (Oxford: Clarendon)

[4] Amit D J, Gutfreund H and Somopolinsky H 1987 Statistical mechanics of neural networks near saturation *Ann. Phys., NY* **173** 30–69

[5] Gardner E and Derrida B 1988 Optimal storage properties of neural network models *J. Phys. A: Math. Gen.* **21** 271–84

[6] Kohring G A 1990 *J. Physique* **51** 145–55

[7] Yamazaki Y, Ochiai M and Holz A 1994 *Proc. 3rd Int. Conf. on Fuzzy Logic* pp 115–6

[8] Yamazaki Y, Kang G and Ochiai M 1998 *Int. J. Intell. Syst.* **13** 503–18

[9] Yamazaki Y, Alyoshin V, Krasilnikova J and Krasilnikov I 1998 *2nd Int. Conf. on Knowledge-Based Intelligent Electronic Systems* pp 230–9

[10] Yamazaki Y, Alyoshin V and Krasilnikova J 1998 *Soft Computing, Multimedia and Image Processing* vol 8 (Albuquerque, NM: TSI Press) pp 557–62

[11] Penny R W, Coolen A C C and Sherrington D 1993 Coupled dynamics of fast spins and slow interactions in neural networks and spin systems *J. Phys. A: Math. Gen.* **26** 3681–95