Effects of neuronal loss in the dynamic model of neural networks

2008 J. Phys. A: Math. Theor. 41 385102

(http://iopscience.iop.org/1751-8121/41/38/385102)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 385102 (10pp) doi:10.1088/1751-8113/41/38/385102

Effects of neuronal loss in the dynamic model of neural networks

B-G Yoon[1], J Choi[2] and M Y Choi[3,4]

[1] Department of Physics, University of Ulsan, Ulsan 680-749, Korea
[2] Department of Physics, Keimyung University, Daegu 704-701, Korea
[3] Department of Physics and Astronomy, Seoul National University, Seoul 151-747, Korea
[4] Asia-Pacific Center for Theoretical Physics, Pohang 790-784, Korea

E-mail: bgyoon@ulsan.ac.kr

Received 22 March 2008, in final form 24 July 2008
Published 18 August 2008
Online at stacks.iop.org/JPhysA/41/385102

Abstract

We study the phase transitions and dynamic behavior of the dynamic model of neural networks, with an emphasis on the effects of neuronal loss due to external stress. In the absence of loss the overall results obtained numerically are found to agree excellently with the theoretical ones. When the external stress is turned on, some neurons may deteriorate and die; such loss of neurons, in general, weakens the memory in the system. As the loss increases beyond a critical value, the order parameter measuring the strength of memory decreases to zero either continuously or discontinuously, namely, the system loses its memory via a second- or a first-order transition, depending on the ratio of the refractory period to the duration of action potential.

PACS numbers: 87.19.lj, 05.10.Gg, 87.18.Sn, 87.19.xr

1. Introduction

Statistical mechanics has contributed much to the understanding of such features as memory, learning, fault tolerance, information storage and its retrieval in terms of collective behaviors of neurons [1–6]. In theoretical approaches, the brain is usually modeled as an ideal network of neurons which are all intact. In the real world, however, a living organism in general suffers from external stress, which may lead the organism into failure. In particular, neurons may deteriorate and die due to stress, which may increase the stress on other neurons, and thus induce death of other neurons. Effects of such loss of neurons may be considered through the use of the dynamic failure model for living organisms, proposed recently for a general system of cells [7]. The model has turned out to display characteristic time evolution, reminiscent of the time progression of degenerative diseases: it tends to resist stress for rather a long time, followed by sudden failure with some fraction of cells possibly surviving.

To describe appropriately the deterioration behavior of a neural network associated with neuronal loss, we need to consider a neural network model with realistic dynamics. Most models usually assume that the dynamics is either totally synchronous or totally asynchronous; this assumption is not realistic in neural networks, in view of the presence of several time scales such as the refractory period, time duration of the action potential and the retardation of signal propagation. On the other hand, the dynamic model for neural networks, proposed by one of the authors, uses a continuous time scale, taking explicitly into account several time scales described above [8]. The dynamic model has been successfully applied to a simple case with a Hebb-type rule [9], presenting the desirable features similar to those in the conventional models. It is thus natural to employ the dynamic model for neural networks, incorporated with the dynamic failure model that reflects the effects of neuronal loss.

In this work, we study the phase transitions and dynamic behavior of the dynamic model of neural networks, incorporated with the dynamic failure model. In the absence of loss, the overall results obtained numerically are in excellent agreement with the theoretical ones. The loss of neurons due to external stress is then shown to weaken the memory present in the system. As the loss increases beyond a critical value, the order parameter measuring the strength of memory decreases to zero either continuously or discontinuously, namely, the system loses the memory via a first- or a second-order transition, depending on the refractory period.

This paper consists of five sections. In section 2, we introduce the dynamic model for neural networks with neuronal loss. Section 3 considers the system in the absence of external stress, and presents the numerical results in comparison with the theoretical ones. The effects of neuronal loss due to external stress are examined in section 4. Finally, a summary is given in section 5.

2. Dynamic model

We consider a network of N neurons, the jth of which is described by the Ising spin $\sigma_j = \pm 1$ (firing/quiescent). The system is also under external stress characterized by load Nf. Each neuron has its own tolerance and endures the stress below the tolerance, remaining alive. The ith neuron may become dead, however, if the tolerance h_i is exceeded. We assign 'living' status variables to these in such a way that $\tau_i = \mp 1$ for the ith neuron alive/dead. The state of the network may then be described by the configuration of all the neurons, i.e., $\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2, \ldots, \sigma_N)$ and $\boldsymbol{\tau} \equiv (\tau_1, \tau_2, \ldots, \tau_N)$. The neurons are interconnected by synaptic junctions of strength $2J_{ij}$ (with $J_{ii} \equiv 0$). The threshold behavior of the ith neuron, namely, whether the neuron fires or not depending on the magnitude of the stimulus, at time t is described by the probability that depends on the difference between the total potential on the ith neuron, V_i, and its threshold value, V_0:

$$V_i - V_0 = \frac{1}{2} \sum_j J_{ij} \sigma'_j (1 - \tau'_j) \equiv E'_i, \tag{1}$$

where it has been assumed as usual that $V_0 = \sum_j J_{ij}$. In equation (1), σ'_j and τ'_j denote the state of the jth neuron and its living status at time $t - t_d$, respectively, with t_d being the delay in interactions including the synaptic delay.

We begin with the conditional probability $p^\sigma(\sigma'_i, t + \delta t | \sigma_i, t; \boldsymbol{\sigma}', t - t_d)$ that ith neuron is in state σ'_i at time $t + \delta t$ given that it is in state σ_i at time t. Since only the neuron in the 'living' state contribute, the conditional probability $p^\sigma(\sigma'_i, t + \delta t | \sigma_i, t; \boldsymbol{\sigma}', t - t_d)$, in the

limit $\delta t \to 0$, can be expressed in terms of the transition rate [8]:

$$p^{\sigma}(\sigma_i'', t+\delta t|\sigma_i, t; \boldsymbol{\sigma}', t-t_d) = \begin{cases} w_i^{\sigma}(\sigma_i; \boldsymbol{\sigma}', t-t_d)\delta t & \text{for} \quad \sigma_i'' = -\sigma_i \\ 1 - w_i^{\sigma}(\sigma_i; \boldsymbol{\sigma}', t-t_d)\delta t & \text{for} \quad \sigma_i'' = \sigma_i, \end{cases} \tag{2}$$

where the transition rate is given by

$$w_i^{\sigma}(\sigma_i; \boldsymbol{\sigma}', t-t_d) = \frac{1-\tau_i}{4t_r}\left[\left(a+\frac{1}{2}\right) + \left(a-\frac{1}{2}\right)\sigma_i + \frac{1-\sigma_i}{2}\tanh\beta E_i'\right], \tag{3}$$

with $a \equiv t_r/t_0$ being the ratio of the refractory period to the action potential duration and the inverse 'temperature' $\beta \equiv 1/T$ measuring the width of the threshold region or the synaptic noise level. In the above equations, $\boldsymbol{\sigma}' \equiv (\sigma_1', \sigma_2', \ldots, \sigma_N')$ represents the configuration of the system at time $t - t_d$, and the dependence of the conditional probabilities on $\boldsymbol{\sigma}'$ is implicit, through E_i'. Here we point out that t_d, t_r, t_0 and accordingly a may be complicated functions of system properties like the average activity and others, which may be incorporated into the model. In this work, we restrict ourselves to the simplest case of these parameters being fixed.

Similarly, the failure dynamics of the network is described by the conditional probability $p^{\tau}(\tau_i = 1, t+\delta t|\tau_i = -1, t; \boldsymbol{\tau}', t-t_D)$ that the ith neuron becomes dead at time $t+\delta t$ given that it is alive at time t [7]. In the limit $\delta t \to 0$, this can be expressed in terms of the transition rate:

$$p^{\tau}(\tau_i'', t+\delta t|\tau_i, t; \boldsymbol{\tau}', t-t_D) = \begin{cases} w_i^{\tau}(\tau_i; \boldsymbol{\tau}', t-t_D)\delta t & \text{for} \quad \tau_i'' = -\tau_i \\ 1 - w_i^{\tau}(\tau_i; \boldsymbol{\tau}', t-t_D)\delta t & \text{for} \quad \tau_i'' = \tau_i, \end{cases} \tag{4}$$

with the transition rate

$$w_i^{\tau}(\tau_i; \boldsymbol{\tau}', t-t_D) = \frac{1-\tau_i}{4t_R}[1 + \tanh\gamma H_i'], \tag{5}$$

where we have assumed that neurons, once they fail, do not regenerate. In equation (4), $\boldsymbol{\tau}' \equiv (\tau_1', \tau_2', \ldots, \tau_N')$ represents the configuration of the system at time $t - t_D$ and $H_i' \equiv f - (h_i/2)(1 - \bar{\tau}')$ (with $\bar{\tau}' \equiv N^{-1}\sum_i \tau_i'$). We have two time scales t_D and t_R: t_D denotes the time delay for the stress redistribution while t_R sets the relaxation time of the failure dynamics. γ^{-1} measures the possible uncertainty (noise level) in the neuron tolerance.

The behavior of the neural network is then governed by the master equation, which describes the evolution of the joint probability $P(\boldsymbol{\tau}, \boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d)$ that the system is in state $\boldsymbol{\tau}'$ at time $t - t_D$, in state $\boldsymbol{\sigma}'$ at time $t - t_d$, and in state $\boldsymbol{\tau}, \boldsymbol{\sigma}$ at time t. One may follow essentially the same procedure as that in [7, 8] to derive the time evolution equation for $P(\boldsymbol{\tau}, \boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d)$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}P(\boldsymbol{\tau}, \boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d) = -\sum_i \big[& w_i^{\tau}(\tau_i; \boldsymbol{\tau}')P(\boldsymbol{\tau}, \boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d) \\ & - w_i^{\tau}(-\tau_i; \boldsymbol{\tau}')P(F_i\boldsymbol{\tau}, \boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d) \\ & + w_i^{\sigma}(\sigma_i; \boldsymbol{\sigma}')P(\boldsymbol{\tau}, \boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d) \\ & - w_i^{\sigma}(-\sigma_i; \boldsymbol{\sigma}')P(\boldsymbol{\tau}, F_i\boldsymbol{\sigma}, t; \boldsymbol{\tau}', t-t_D, \boldsymbol{\sigma}', t-t_d)\big], \end{aligned} \tag{6}$$

where we have used the abbreviations $w_i^{\tau}(\tau_i; \boldsymbol{\tau}') \equiv w_i^{\tau}(\tau_i; \boldsymbol{\tau}', t-t_D)$ with $F_i\boldsymbol{\tau} \equiv (\tau_1, \ldots, \tau_{i-1}, -\tau_i, \tau_{i+1}, \ldots, \tau_N)$, and $w_i^{\sigma}(\sigma_i; \boldsymbol{\sigma}') \equiv w_i^{\sigma}(-\sigma_i; \boldsymbol{\sigma}', t-t_d)$ with $F_i\boldsymbol{\sigma} \equiv (\sigma_1, \ldots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \ldots, \sigma_N)$.

Then equations describing the time evolution of relevant physical quantities, in general, assume the form of retarded differential equations. For instance, the 'living' status variable,

$m_k(t) \equiv \langle\tau_k\rangle_t = \sum_{\tau,\tau'}\sum_{\sigma,\sigma'}\tau_k P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)$ may be obtained from equation (6), by multiplying τ_k and summing over all configurations:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}m_k = -\sum_{\tau,\tau'}\sum_{\sigma,\sigma'}\tau_k\big[&w_k^{\tau}(\tau_k;\boldsymbol{\tau}')P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &- w_k^{\tau}(-\tau_k;\boldsymbol{\tau}')P(F_k\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &+ w_k^{\sigma}(\sigma_k;\boldsymbol{\sigma}')P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &- w_k^{\sigma}(-\sigma_k;\boldsymbol{\sigma}')P(\boldsymbol{\tau},F_k\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\big]\\ = -2&\big\langle\tau_k w_k^{\tau}(\tau_k;\boldsymbol{\tau})\big\rangle_t,\end{aligned} \tag{7}$$

where it has been utilized that

$$\begin{aligned}\sum_{\tau,\tau'}\sum_{\sigma,\sigma'}\big[&w_k^{\sigma}(\sigma_k;\boldsymbol{\sigma}')P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &- w_k^{\sigma}(-\sigma_k;\boldsymbol{\sigma}')P(\boldsymbol{\tau},F_k\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\big]=0.\end{aligned} \tag{8}$$

Evaluation of the average $\langle\tau_k w_k^{\tau}(\tau_k;\boldsymbol{\tau})\rangle_t$ leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}m_k(t) = \frac{1}{t_R}\left[\frac{1-m_k}{2} + \left\langle\frac{1-\tau_k}{2}\tanh\gamma H_k'\right\rangle_t\right]. \tag{9}$$

In particular, the 'effective' activity of the kth neuron, which reflects the activity of only a living one,

$$s_k(t) \equiv \left\langle\frac{1-\tau_k}{2}\sigma_k\right\rangle_t \equiv \sum_{\tau,\tau'}\sum_{\sigma,\sigma'}\frac{1-\tau_k}{2}\sigma_k P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d), \tag{10}$$

can be determined from

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}s_k = -\sum_{\tau,\tau'}\sum_{\sigma,\sigma'}\sigma_k\frac{1-\tau_k}{2}\big[&w_k^{\tau}(\tau_k;\boldsymbol{\tau}')P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &- w_k^{\tau}(-\tau_k;\boldsymbol{\tau}')P(F_k\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &+ w_k^{\sigma}(\sigma_k;\boldsymbol{\sigma}')P(\boldsymbol{\tau},\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\\ &- w_k^{\sigma}(-\sigma_k;\boldsymbol{\sigma}')P(\boldsymbol{\tau},F_k\boldsymbol{\sigma},t;\boldsymbol{\tau}',t-t_D,\boldsymbol{\sigma}',t-t_d)\big]\\ = \big\langle\tau_k\sigma_k w_k^{\tau}(\tau_k;\boldsymbol{\tau})\big\rangle_t &- \big\langle(1-\tau_k)\sigma_k w_k^{\sigma}(\tau_k;\boldsymbol{\tau})\big\rangle_t.\end{aligned} \tag{11}$$

Evaluation of the average $\big\langle\tau_k\sigma_k w_k^{\tau}(\tau_k;\boldsymbol{\tau})\big\rangle_t$ and $\big\langle(1-\tau_k)\sigma_k w_k^{\sigma}(\tau_k;\boldsymbol{\tau})\big\rangle_t$ leads to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}s_k = -\frac{1}{2t_R}\left[s_k + \left\langle\frac{1-\tau_k}{2}\sigma_k\tanh\gamma H_k'\right\rangle_t\right] + \frac{1}{t_r}\Bigg[&\left(\frac{1}{2}-a\right)\frac{1-m_k}{2} - \left(\frac{1}{2}+a\right)s_k\\ &+ \frac{1}{2}\left\langle\frac{1-\tau_k}{2}\tanh\beta E_k'\right\rangle_t - \frac{1}{2}\left\langle\frac{1-\tau_k}{2}\sigma_k\tanh\beta E_k'\right\rangle_t\Bigg].\end{aligned} \tag{12}$$

We now consider the Hebb-type synaptic couplings

$$J_{ij} = \begin{cases} N^{-1}\sum_{\mu=1}^{p}\xi_i^{\mu}\xi_j^{\mu}, & i\neq j,\\ 0, & i=j,\end{cases} \tag{13}$$

which corresponds to the situation that p patterns $\{\xi_i^{\mu}\}$ $(\mu = 1, 2, \ldots, p)$ are learned. Since this assignment implies that J_{ij} is of infinite range, the mean-field approximation is expected to be correct. One may then replace E_k' in equation (12) by its average $\langle E_k'\rangle = \sum_{\mu}\xi_k^{\mu}q^{\mu}$ with the 'order parameter' $q^{\mu} \equiv N^{-1}\sum_j \xi_j^{\mu}s_j$ describing the overlap between the network state and the memory. Further, in the case of the simplest global load sharing, we may also replace the local field H_k' by its average $\langle H_k'\rangle = f - (h_k/2)[1 - m(t - t_D)]$ with

$m(t-t_D) \equiv N^{-1}\sum_k m_k(t-t_D)$. Defining $x_k \equiv (1-m_k)/2$ and the average number of intact neurons $\bar{x} \equiv N^{-1}\sum_k x_k = (1-\bar{m})/2$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}x_k(t) = -\frac{1}{2t_R}\left[x_k(t) + x_k(t)\tanh\gamma[f - h_k\bar{x}(t-t_D)]\right]. \tag{14}$$

Equation (12) then reads

$$\frac{\mathrm{d}}{\mathrm{d}t}s_k = -\frac{1}{2t_R}[s_k + s_k\tanh\gamma[f - h_k\bar{x}(t-t_D)]] + \frac{1}{t_r}\left[\left(\frac{1}{2}-a\right)x_k - \left(\frac{1}{2}+a\right)s_k + \frac{1}{2}x_k\tanh\beta\boldsymbol{\xi}_k\cdot\mathbf{q}(t-t_d) - \frac{1}{2}s_k\tanh\beta\boldsymbol{\xi}_k\cdot\mathbf{q}(t-t_d)\right], \tag{15}$$

where the vector notation $\boldsymbol{\xi}_k\cdot\mathbf{q} \equiv \sum_\mu \xi_k^\mu q^\mu$ has been introduced for simplicity.

Multiplying equation (15) by $N^{-1}\boldsymbol{\xi}_k$ and summing over k, we obtain the equation for the order parameter:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q} = -\frac{1}{2t_R}\left[\mathbf{q} + \frac{1}{N}\sum_k \boldsymbol{\xi}_k s_k\tanh\gamma[f - h_k\bar{x}(t-t_D)]\right] + \frac{1}{t_r}\left[-\left(\frac{1}{2}+a\right)\mathbf{q} + \frac{1}{2N}\sum_k \boldsymbol{\xi}_k x_k[1-2a+\tanh\beta\boldsymbol{\xi}_k\cdot\mathbf{q}(t-t_d)] - \frac{1}{2N}\sum_k \boldsymbol{\xi}_k s_k\tanh\beta\boldsymbol{\xi}_k\cdot\mathbf{q}(t-t_d)\right]. \tag{16}$$

Thus, the time evolution of the system is described by three coupled retarded differential equations (14)–(16).

It is obvious that equation (16) has a trivial solution $\mathbf{q} = 0$ together with $x_k = 0$, which corresponds to the disordered state with no memory. We now examine the stationary solutions of the system, which are obtained by setting the time derivatives equal to zero. From equation (14), one finds $x_k = 0$ or

$$\tanh\gamma[f - h_k\bar{x}(t-t_D)] = -1, \tag{17}$$

which leads to the average fraction of living neurons in the noiseless limit ($\gamma^{-1} = 0$):

$$\bar{x} = \int \mathrm{d}h\, g(h)x = \int \mathrm{d}h\,\theta(h\bar{x} - f)g(h), \tag{18}$$

with $g(h)$ being the distribution function of the tolerance $\{h_k\}$. At $\gamma^{-1} \neq 0$, we have $x_k = 0$, giving $\bar{x} = 0$ as the only possible solution. Note that equation (17) implies that the terms containing t_R in equations (14)–(16) vanish. Solving s_k in terms of x_k in equation (15) and substituting into equation (16), we obtain

$$\mathbf{q} = \frac{4a}{1+2a}\left\langle\!\!\left\langle \frac{x\boldsymbol{\xi}\tanh\beta\boldsymbol{\xi}\cdot\mathbf{q}}{1+2a+\tanh\beta\boldsymbol{\xi}\cdot\mathbf{q}}\right\rangle\!\!\right\rangle, \tag{19}$$

where $\langle\!\langle\cdots\rangle\!\rangle$ stands for the average taken with respect to the distribution of $\{\xi_k^\mu\}$ and $\{h_k\}$. In deriving equation (19), we have replaced the average over the neurons by the average over the distributions of memories and of tolerances: $N^{-1}\sum_k f(\boldsymbol{\xi}_k, x_k) = \langle\!\langle f(\boldsymbol{\xi}, x)\rangle\!\rangle$. This self-averaging property should be valid for $N \to \infty$ and p fixed if the memories are essentially random.

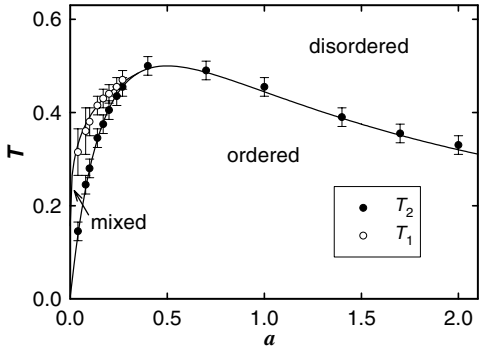

Figure 1. Phase diagram for the Mattis state, which shows the disordered state, ordered state and mixed state, depending on the temperature T and the parameter a, and the ratio of the refractory period to the duration of the action potential. Solid lines are obtained from equation (21) while symbols are from equation (15) with $\bar{x} = 1$, both of which represent the phase boundaries.

3. Behavior in the absence of stress

In this work, we focus on the Mattis-state solution of the form $\mathbf{q} = (q, 0, \ldots, 0)$, which is fully correlated with just one of the quenched memories. For simplicity, ξ_i^μ may be taken to be quenched random variables, assuming $+1$ and -1 with equal probabilities:

$$P(\boldsymbol{\xi}) = \prod_\mu \prod_k p(\xi_k^\mu),$$
$$p(\xi^\mu) = \frac{1}{2}\delta(\xi^\mu - 1) + \frac{1}{2}\delta(\xi^\mu + 1). \tag{20}$$

There are $2p$ equivalent solutions for a Mattis state, and for such solutions it is straightforward to take the average in equation (19) over the distribution given by equation (20). The average over $\{h_k\}$ replaces x by its stationary value $\bar{x}$. Equation (19) for the ideal Mattis state then reduces to the form

$$q = \bar{x}\frac{4a\tanh\beta q}{(1+2a)^2 - \tanh^2\beta q} \equiv g(q), \tag{21}$$

which may be solved numerically to obtain the order parameter as a function of other parameters. For comparison, we reproduce the phase diagram for the system of $\bar{x} = 1$ [8], plotted as solid lines in figure 1. The system is shown to undergo a transition from the disordered state ($q = 0$) to the ordered one with finite q as T is lowered below the transition temperature. When the ratio a is larger than the critical value $a_c \equiv (\sqrt{3} - 1)/2$, the transition is continuous at transition temperature T_2. Below a_c, there appears another phase between T_2 and another temperature T_1 ($>T_2$), in which the ordered state and the disordered state coexist depending on the initial configuration of the system. The transition at $T = T_1$ is discontinuous, between the disordered state to the mixed one. As T is lowered below T_2, the system becomes ordered via the continuous transition.

We also obtain the transition temperatures by means of integrating numerically equation (15) as follows. We consider systems of $N = 4096$ neurons with p memory patterns, which are generated according to equation (20) at each run. The initial values of m_k's are also chosen randomly ($-1 \leqslant m_k \leqslant 1$) at each run, which, together with p patterns, comprise one configuration. Specifically, we have used 10^2 to 10^5 different configurations for each value of

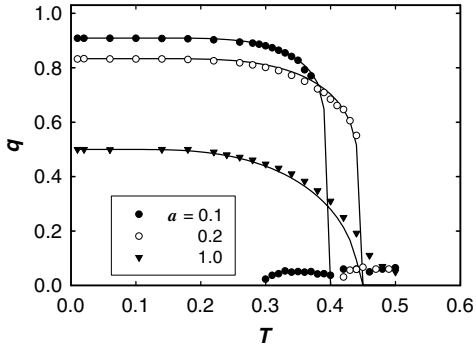

Figure 2. The Mattis-state order parameter q versus the temperature T for $\bar{x} = 1$. Solid lines are obtained from equation (21) and symbols from equation (15). Full and open circles represent the data for $a = 0.1$ and 0.2, respectively, while back triangles correspond to $a = 1.0$.

(a, T), and set the refractory period $t_r = 5$, delay time $t_d = 1$ and time step $\Delta t = 0.5$. These parameter values have been varied, only to give no appreciable difference except for the time scale. In the stationary state, the first term in the right-hand side of equation (15) vanishes and only the second term needs to be integrated.

Raising the temperature stepwise in the increment $\Delta T = 0.01$, we record at each temperature all configurations, which may correspond to the disordered, Mattis or other ordered state. Among ordered states of the system, only Mattis states are taken, from which the transition temperatures are estimated. Here it is rather arduous to determine accurately the transition temperature, for the following reason. At temperatures near T_2, realization of a Mattis state is of extremely small probability, especially as the number p of patterns is increased. The same is true for the realization of a mixed state for $a < a_c$ at temperatures near T_1. Fortunately, the data appear largely independent on the choice of p, which allows one to use small values of p. We thus use the values between 2 and 10, to obtain the transition temperatures T_1 and T_2. Figure 1 exhibits the resulting data, represented by open (T_1) and filled (T_2) circles; good agreement with the analytical results is observed.

In figure 2, we plot the order parameter q as a function of temperature T, to disclose the nature of the phase transition. Whereas solid lines are obtained from equation (21), symbols represent data from integrating equation (15). Full and open circles correspond to data for $a = 0.1$ and 0.2, respectively, and back triangles to data for $a = 1.0$. When $a < a_c$, the data in circles show that the transition is discontinuous and the order parameter can take two different values, thus manifesting the mixed state, at temperatures between T_1 and T_2. As a is increased beyond a_c, on the other hand, the transition becomes continuous, without the mixed state (see the data for $a = 1.0$).

In addition, we have also performed dynamic Monte Carlo (MC) simulations of the system, using the transition probabilities in equations (2) and (4). The results agree well with those obtained by integrating equation (15) with $\bar{x} = 1$. Thus analytical predictions such as the existence of the mixed state for $a < a_c$ and the nature of the transition are confirmed by both numerical integration and simulations.

4. Effects of neuronal loss

When the stress is turned on, the fraction $\bar{x}$ of intact neurons will decrease. In calculation with the deterioration taken into account, we take the tolerance distribution $g(h)$ to be Gaussian

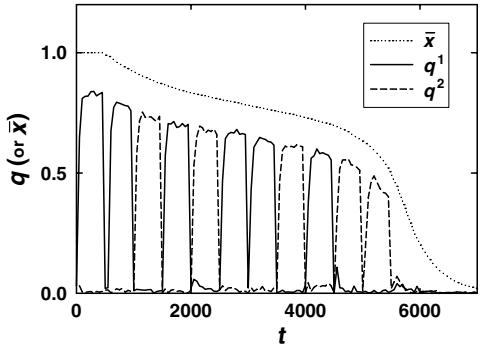

Figure 3. Evolution of the Mattis-state order parameter q with time t, obtained from Monte Carlo simulations based on equations (2) and (4). The solid line and the dashed one correspond to two different memory patterns; also shown is the evolution of the fraction $\bar{x}$ of intact neurons (plotted by the dotted line). The tolerance distribution is taken to be Gaussian with unit mean and variance $\sigma = 0.3$, the stress $f = 0.60$. The time scale of the deterioration process is taken to be 100 times that of the memory function process.

with unit mean and variance $\sigma = 0.3$ or 0.4, and set the relaxation time $t_R = 50$ and the delay time for the stress redistribution $t_D = 100$.

Monte Carlo simulations reveal that the Mattis-state order parameter q decreases as the deterioration goes on; finally, the system loses memory completely at some value of $\bar{x}$, which depends on the parameters of the system. As an example, figure 3 displays a typical result of the deterioration process of the system with two patterns ($p = 2$) for $a = 0.2$ and $T = 0.2$ under stress $f = 0.60$ turned on at time $t = 500$. Beginning in the disordered state ($q^1 = q^2 = 0$ at $t = 0$), the system rapidly evolves into the ordered state characterized by $q^1 \neq 0$, namely, the system retrieves quickly the memory ξ^1, which is expected at $T < T_2$. To probe such retrieval capability in the presence of the deterioration, we erase memory by resetting $q = 0$ at time $t = 500$ and periodically thereafter (with the period of 500) while keeping the stress turned on. It is observed that under such memory resetting, the system still evolves rapidly into an ordered state, retrieving one pattern of memory, until the fraction $\bar{x}$ of intact neurons, decreasing from unity, reaches some threshold value. As the fraction decreases below this value, the memory ceases to be retrieved.

It is obvious that q in general reduces as the fraction $\bar{x}$ decreases or the neuronal loss $(1 - \bar{x})$ increases. This may be demonstrated through the use of equation (21), as shown by solid lines in figure 4 for $a = 0.1$ and at two temperatures. Data represented by symbols have been obtained from the integration of equation (15). Here note that the time scale of the deterioration process is much longer than that of the memory retrieval process. The state of the system would thus be practically quasi-stationary except during the final breakdown.

We also show the results for $a = 0.5$ in figure 5, where the symbols represent data obtained from MC simulations rather than from numerical integration. In the simulations, two configurations have been used: variance $\sigma = 0.4$ for the tolerance distribution, stress $f = 0.61$, and temperature $T = 0.1$ and 0.3. Larger discrepancy is observed between the analytical results (solid lines) and simulation results near the threshold region, where the system evolves rapidly, not in a quasi-stationary state. We have also integrated numerically equation (15), similarly to the case of figure 4, to obtain results in excellent agreement with the analytical ones (data not shown).

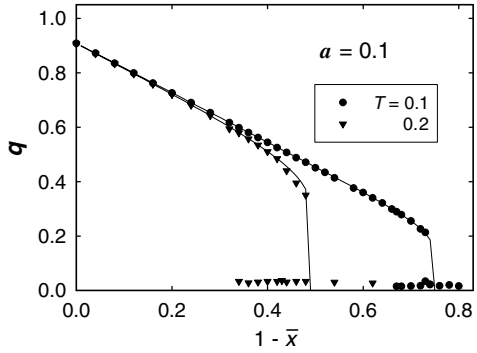

Figure 4. The Mattis-state order parameter q versus the neuronal loss $(1 - \bar{x})$ for $a = 0.1$ at temperatures $T = 0.1$ and 0.2. Solid lines are obtained from equation (21) and symbols represent data from equation (15). Discontinuous changes of the order parameter indicate the presence of first-order transitions.

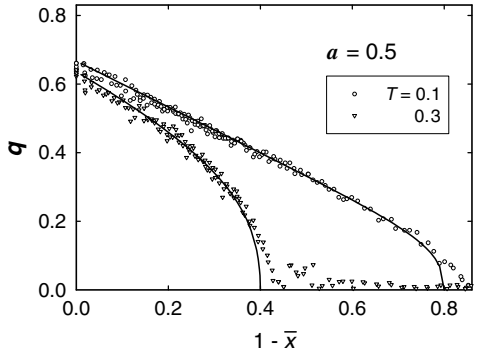

Figure 5. The Mattis-state order parameter q versus the neuronal loss $(1 - \bar{x})$ for $a = 0.1$ at temperature $T = 0.1$ and 0.3. Solid lines are obtained from equation (21) while symbols represent data from Monte Carlo simulations, similar to those used in obtaining the data in figure 2. Second-order transitions are observed.

In figures 4 and 5, the order parameter reduces as the neuronal loss increases. Then the system loses memory completely when the loss becomes larger than a threshold value. As the temperature is increased, this behavior is enhanced, i.e., q decreases faster. The transition is discontinuous for $a < a_c$, as shown in figure 4, and continuous for $a > a_c$ as shown in figure 5, which is similar to the transition with the temperature. When a is small, there also exists a range of the neuronal loss in which the mixed phase appears; this is manifested by the presence of two different values of q in figure 4. The system with a small ratio of the refractory period to the action potential duration may thus be either successful or failing in retrieving the memory, even in the simplest mean-field-like theory. In real systems, these results may be related to, for example, dementia or Alzheimer's disease [10], in which the system loses memory as the neuronal loss proceeds. It is thus of interest to see whether the refractory period or the action potential duration alters in the neuronal cells of a patient suffering from the disease.

5. Summary

We study, both theoretically and numerically, the effects of neuronal loss on the phase transition of the neural network. In the absence of loss, both numerical integration and simulations give results, which agree well with those obtained analytically. When the ratio a of the refractory period to the action potential duration is smaller than the critical value, the transition is discontinuous, and there exists a mixed state at intermediate temperatures. As a grows beyond the critical value, the transition turns into a continuous one, without a mixed state.

When the neuronal loss is introduced, the Mattis-state order parameter in general reduces with the loss. Above some threshold value, the system loses memory completely, as expected. There occurs either a first- or a second-order transition as the neuronal loss increases, depending on the parameter a; this corresponds to the memory retrieved abruptly or gradually.

Acknowledgments

M Y C thanks H E Stanley for hospitality during his stay at Boston University, where part of this work was accomplished. This work was supported in part by the 2007 Research Fund of the University of Ulsan.

References

[1] Little W A 1974 *Math. Biosci.* **19** 101
Little W A and Shaw G A 1978 *Math. Biosci.* **39** 281
[2] Hopfield J J 1982 *Proc. Natl. Acad. Sci. USA* **79** 2554
[3] Peretto P 1984 *Biol. Cybern.* **50** 51
Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev.* A **32** 1007
Fontanari J F and Köberle R 1987 *Phys. Rev.* A **36** 2475
[4] Sompolinsky H and Kanter I 1986 *Phys. Rev. Lett* **57** 2861
Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett.* **4** 167
Crisanti A and Sompolinsky H 1988 *Phys. Rev.* A **37** 4865
[5] Meir R and Domany E 1988 *Phys. Rev.* A **37** 2660
[6] Watkin T L H, Rau A and Biehl M 1993 *Rev. Mod. Phys.* **65** 499
Hopfield J J 1999 *Rev. Mod. Phys.* **71** S431
[7] Choi M Y, Choi J and Yoon B-G 2004 *Europhys. Lett.* **66** 62
Choi J, Choi M Y and Yoon B-G 2005 *Europhys. Lett.* **71** 501
Yoon B-G, Choi J and Choi M Y 2005 *J. Korean Phys. Soc.* **47** 1053
Yoon B-G, Choi J, Choi M Y and Fortin J-Y 2006 *Phys. Rev.* E **73** 031905
[8] Choi M Y 1988 *Phys. Rev. Lett.* **61** 2809
[9] Hebb D O 1949 *The Organisation of Behaviour* (New York: Wiley)
[10] See, e.g., Terry R D, Katzman R and Bick K (ed) 1994 *Alzheimer's Disease* (New York: Raven)
Vincent I, Jicha G, Rosado M and Dickson D W 1997 *J. Neurosci.* **17** 3588 and references therein