Neural network learning dynamics in a path integral framework

J. Balakrishnan[a]

Department of High Voltage Engineering, Indian Institute of Science, Bangalore, 560 012, India

Received 19 November 1999

Abstract. A path-integral formalism is proposed for studying the dynamical evolution in time of patterns in an artificial neural network in the presence of noise. An effective cost function is constructed which determines the unique global minimum of the neural network system. The perturbative method discussed also provides a way for determining the storage capacity of the network.

PACS. 02.90.+p Other topics in mathematical methods in physics – 05.90.+m Other topics in statistical physics, thermodynamics, and nonlinear dynamical systems – 87.10.+e General theory and mathematical aspects

1 Introduction

It has been of interest since long to understand the mechanism of learning and memory in biological systems and machines. Studies of associative memory have sought to model the process of pattern recognition and recall using specific cost functions.

The similarity of the McCulloch-Pitts neural network to the Ising spin system has enabled statistical physics approaches [1–4] to be used to get information like the storage capacity of the network. In the "space of interactions" approach of Gardner *et al.* [2,3], an imbedding condition is postulated and the energy function counts the number of weakly-imbedded pattern spins which have stability less than a specified value. Though this approach does a systematic study of a neural network configuration at any given instance of time, it does not address the question of time evolution of the network configuration. Some years ago, Hertz *et al.* [5,6] studied the dynamics of learning in a single-layer neural network using a Langevin equation for the evolution in time of the synaptic efficacies. In these papers the authors have investigated the role of noise in learning and studied the possible phase transitions in the learning process.

In our work we have taken such a viewpoint, of looking at the learning process as a non-equilibrium stochastic process, as our starting point for constructing a path-integral framework for studying neural network dynamics.

The problem of neural networks getting trapped into spurious states or local minima is well-known and a method for directly getting to the global minimum of the network is highly desirable. Of much more interest is a systematic theory which gives a framework for determining the global minimum of the neural network model, independent of the choice of the cost function.

[a] e-mail: `janaki@serc.iisc.ernet.in`

An attempt has been made in this work to achieve this through the path-integral framework using concepts from quantum and statistical field theory — we have considered a perceptron only for the sake of simplicity.

In this framework it is seen that the patterns in the network settle from non-equilibrium states into certain attractor states which correspond to those with lowest energy at equilibrium, or for large values of the time. We construct an "effective cost function" for any cost function one starts with, and discuss how the global minimum for a neural network can be determined and how one can calculate the storage capacity of the network with this construction.

2 Langevin dynamics in a path-integral approach

Langevin dynamics has been applied to analyse disordered systems, in particular, to spin glasses and to retrieval processes in attractor neural network models with fixed weights, by a number of authors [7]. In our work, we look at learning dynamics from a slightly different viewpoint — through a path-integral framework, and with a dynamical evolution in time for the synaptic efficacies.

As in [5,6], we view the problem of learning in neural networks as a stochastic process and for simplicity, we look at the perceptron only, with one layer of connections. We postulate a stochastic Langevin dynamics for the evolution in time of the synaptic efficacies w_{ij}

$$\frac{\partial \omega_i(\tau)}{\partial \tau} = -\gamma T \frac{\delta E(\omega_i)}{\delta \omega_i(\tau)} + \eta_i(\tau) \qquad (1)$$

where the input index i takes values from 1 to N and we have omitted the output index which can be treated separately. $\eta_i(t)$ stands for a random white noise source,

T is the noise level, E is the cost function and the parameter γ describes the learning rate. The system can alternatively be thought to be coupled to a heat reservoir at temperature T (which represents the noise level) and evolves in time t until it reaches an equilibrium configuration at $t \to \infty$.

Here, and in the following, the space label has been suppressed and the index i should be interpreted to include the space variable also. We assume that the noise sources are Gaussian with the correlations:

$$\langle \eta_i(\tau) \rangle = 0$$
$$\langle \eta_i(\tau) \eta_j(\tau') \rangle = 2\gamma T \delta_{ij} \delta(\tau - \tau'). \tag{2}$$

The values of the synaptic efficacies at any instant of time are determined by solving (1) and the correlations between them can be calculated using (2).

We wish to construct a partition function and a path integral framework for this simplest type of neural network which incorporates the time evolution of the synaptic strengths. Pattern recall and recognition takes place best when the Hamming distance between the target ζ_i^μ for pattern $\mu = 1, \ldots, p$ and the input pattern ξ_i^ν is minimal for $\mu = \nu$. This means that the synaptic strengths ω_{ij} change in such a way as to descend the cost function surface.

In an artificial neural network, the problem of the system getting trapped in its various local minima or spurious states is a familiar one. Here, prior knowledge of the global minimum for the system and the synaptic efficacies corresponding to it would be very useful as it would save a great deal of effort and computer time and enable greater efficiency in solving pattern recognition and associative memory problems.

We adapt the procedure which was elucidated by Gozzi [8] in a different context, to the neural network system and write down a partition function for this system evolving in time through the Langevin dynamics of (1):

$$Z[J] = \mathcal{N} \prod_{i,\mu} \int_0^\tau \mathrm{D}\omega_i \mathrm{D}\eta_i \, \mathrm{e}^{-\frac{1}{\gamma T} \int J_i(\tau') \omega_i(\tau') \mathrm{d}\tau'}$$
$$\times P(\omega(0)) \delta(\omega_i - \omega_{i\eta}) \mathrm{e}^{-\int \frac{\eta_i^2}{4\gamma T} \mathrm{d}\tau'} \tag{3}$$

where $\mathcal{N}$ is a normalization constant, and we have introduced an external source $J_i(t)$ which probes the fluctuations of the ergodic ensemble. At the end of the calculations, in the thermodynamic limit $N \to \infty$, J_i would be set to zero. $\omega_{i\eta}$ is the solution of the Langevin equation (1) with the initial probability distribution $P(\omega(0))$.

Using some algebraic manipulations we now rewrite this partition function in such a manner that the dynamical evolution of the network configuration in time becomes more apparent and it is seen that for large values of the time, the patterns evolve into the configuration corresponding to the ground state energy of the system.

From (1), we can write:

$$\delta(\omega_i - \omega_{i\eta}) = \delta\left(\dot{\omega}_i + \gamma T \frac{\delta E}{\delta \omega_i} - \eta_i\right) \left\|\frac{\delta \eta_i}{\delta \omega_i}\right\| \tag{4}$$

where the Jacobian $\left\|\frac{\delta \eta}{\delta \omega}\right\|$ of the transformation $\eta_i \to \omega_i$ can be written as:

$$\left\|\frac{\delta \eta_i}{\delta \omega_j}\right\| = \det\left[\left(\delta_{ij} \partial_\tau + \gamma T \frac{\delta^2 E}{\delta \omega_i(\tau) \delta \omega_j(\tau')}\right) \delta(\tau - \tau')\right]$$
$$= \exp\left[\mathrm{tr} \ln \partial_\tau \left(\delta_{ij} \delta(\tau - \tau') + \partial_{\tau'}^{-1} \gamma T \frac{\delta^2 E}{\delta \omega_i(\tau) \delta \omega_j(\tau')}\right)\right] \tag{5}$$

and $\partial_{\tau'}^{-1}$ satisfies:

$$\partial_\tau G(\tau - \tau') = \delta(\tau - \tau'). \tag{6}$$

Equation (4) then reduces to:

$$\left\|\frac{\delta \eta_i}{\delta \omega_i}\right\| = \exp\left\{\mathrm{tr}\left[\ln \partial_\tau + \ln\left(\delta(\tau - \tau') + G_{ij}(\tau - \tau') \gamma T \frac{\delta^2 E}{\delta \omega_j(\tau) \delta \omega_i(\tau')}\right)\right]\right\}. \tag{7}$$

Since we are primarily interested in the dynamical evolution of the patterns of the neural network forward in time, the Green's function $G(\tau - \tau')$ must satisfy

$$G(\tau - \tau') = \theta(\tau - \tau'). \tag{8}$$

Substituting for G from (7) into (6), and expanding out the logarithm we obtain

$$\left\|\frac{\delta \eta_i}{\delta \omega_i}\right\| = \mathrm{e}^{\gamma T \int_0^\tau \mathrm{d}\tau' \theta(0) \frac{\delta^2 E}{\delta \omega_i(\tau')^2}} \tag{9}$$

where an overall factor of $\mathrm{tr} \ln \partial_\tau$ in the exponential has been absorbed in the normalization. We choose to work with the mid-point prescription $\theta(0) = 1/2$ which leads to

$$\left\|\frac{\delta \eta_i}{\delta \omega_i}\right\| = \mathrm{e}^{\frac{\gamma T}{2} \int_0^\tau \mathrm{d}\tau' \frac{\delta^2 E}{\delta \omega_i(\tau')^2}}. \tag{10}$$

From (1), (3) and (10) the partition function can be rewritten as

$$Z[J] = \mathcal{N} \prod_{i,\mu} \int \mathrm{D}\omega_i \, P(\omega(0)) \mathrm{e}^{-\frac{1}{\gamma T} \int_0^\tau J_i(\tau') \omega_i(\tau') \mathrm{d}\tau'}$$
$$\times \mathrm{e}^{\frac{\gamma T}{2} \int_0^\tau \mathrm{d}\tau' \frac{\delta^2 E}{\delta \omega_i^2}} \mathrm{e}^{-\frac{1}{4\gamma T} \int_0^\tau \mathrm{d}\tau' (\dot{\omega}_i + \gamma T \frac{\delta E}{\delta \omega_i})^2}$$
$$= \mathcal{N} \prod_{i,\mu} \int \mathrm{D}\omega_i \, P(\omega(0))$$
$$\times \mathrm{e}^{-\int_0^\tau \mathrm{d}\tau' \{\frac{1}{\gamma T} J_i(\tau') \omega_i(\tau') + \frac{1}{4\gamma T}[\dot{\omega}_i^2 + (\gamma T)^2(\frac{\delta E}{\delta \omega_i})^2 - 2(\gamma T)^2 \frac{\delta^2 E}{\delta \omega_i^2}]\}}$$
$$\times \mathrm{e}^{-\frac{1}{2} E(\tau)} \mathrm{e}^{\frac{1}{2} E(0)}. \tag{11}$$

The N-point correlations between the synaptic efficacies can be calculated by taking the N th functional derivative of $Z[J]$ with respect to J. After making a change of scale of the time variable τ as: $\tau' \to \tau'' = 2\tau'$ we find that the partition function can be rewritten as

$$Z[J] = \mathcal{N} \prod_{i,\mu} \int \mathrm{D}\omega_i P(\omega(0))$$
$$\times \mathrm{e}^{-\frac{1}{\gamma T} \int_0^{2\tau} \mathcal{L}^{FP} \mathrm{d}\tau' - \int_0^{2\tau} \mathrm{d}\tau'' \frac{1}{\gamma T} J_i(\tau'') \omega_i(\tau'')} \mathrm{e}^{-\frac{E(2\tau)}{2}} \mathrm{e}^{\frac{E(0)}{2}} \tag{12}$$

where we have defined a Fokker-Planck Lagrangian as:

$$\mathcal{L}^{\mathrm{FP}} = \frac{1}{2}\left(\frac{\mathrm{d}\omega_i}{\mathrm{d}\tau}\right)^2 + \frac{(\gamma T)^2}{8}\left(\frac{\delta E}{\delta\omega_i}\right)^2 - \frac{\gamma T}{4}\frac{\delta^2 E}{\delta\omega_i^2}. \quad (13)$$

It may be observed that this can be derived from a Fokker-Planck Hamiltonian H^{FP} which obeys the heat equation

$$\frac{\partial\Psi}{\partial\tau} = -2H^{\mathrm{FP}}\Psi \quad (14)$$

where

$$\Psi = \mathrm{e}^{\frac{E(\omega)}{2}}P(\omega,\tau) \quad (15)$$

and

$$H^{\mathrm{FP}} = \gamma T\left[-\frac{1}{2}\frac{\delta^2}{\delta\omega_i^2} + \frac{1}{8}\left(\frac{\delta E}{\delta\omega_i}\right)^2 - \frac{1}{4}\frac{\delta^2 E}{\delta\omega_i^2}\right]. \quad (16)$$

It is possible to find a series solution to (14):

$$\Psi(\omega,\tau) = \sum_n c_n\psi_n\mathrm{e}^{-2E_n\tau} \quad (17)$$

where c_n are normalizing constants and $E_n \geq 0$ are the energy eigenvalues of the operator equation:

$$H^{\mathrm{FP}}\psi_n = E_n\psi_n. \quad (18)$$

It is easy to see that H^{FP} is a positive semi-definite operator with its ground state $E_0 = 0$ defined by $\psi_0 = \mathrm{e}^{-\frac{E(\omega)}{2}}$. In the equilibrium limit $t \to \infty$, only the ground state configuration ψ_0 contributes to $P(\omega,\tau)$ and we have

$$\lim_{t\to\infty} P(\omega,\tau) = c_0\mathrm{e}^{-E(\omega)}. \quad (19)$$

Interestingly, following Gozzi [9], it is also possible to define a canonical momentum $\hat{\Pi}$ conjugate to the variable representing the synaptic efficacy, from the Fokker-Planck Lagrangian:

$$\hat{\Pi}_i = \frac{\delta\mathcal{L}^{\mathrm{FP}}}{\delta\dot{\omega}_i} = \frac{1}{2\gamma T}\left(\dot{\omega}_i + \gamma T\frac{\delta E}{\delta\omega_i}\right) \quad (20)$$

so that the partition function can be written in the form of the Gibbs average of equilibrium statistical mechanics:

$$Z[J] = \\ \mathcal{N}\prod_{i,\mu}\int\mathcal{D}\omega_i(0)\mathcal{D}\hat{\Pi}_i(0)\mathrm{e}^{-\int\{\frac{1}{2}\hat{\Pi}_i^2 + E(\omega) + J_i(\tau')\omega_i(\tau')\}\mathrm{d}\tau'}. \quad (21)$$

The advantage of writing Z in this form is that the integration measure and Z are independent of τ.

The correlations between the synaptic strengths can be calculated either using equations (2) within the Langevin

approach, or using the equivalent Fokker-Planck equations. Since the noise-sources are delta-correlated, the correlations between the synaptic efficacies would be stationary for a proper choice of the initial probability $P(\omega(0))$:

$$\langle\omega_\eta(\tau_1)\ldots\omega_\eta(\tau_l)\rangle_{\eta,P(\omega(0))} = \\ \delta(\tau_1 - \tau_2, \tau_2 - \tau_3, \ldots, \tau_l - \tau_{l-1}) \quad (22)$$

where $\langle.\rangle_{\eta,P(\omega(0))}$ denotes average over both η and $\omega(0)$. Since the correlations depend only on the time differences, these would remain invariant under a uniform translation in the time:

$$\langle\omega_\eta(\tau_1)\ldots\omega_\eta(\tau_l)\rangle_{\eta,P(\omega(0))} = \\ \langle\omega_\eta(\tau_1 + t)\ldots\omega_\eta(\tau_l + t)\rangle_{\eta,P(\omega(0))}. \quad (23)$$

Setting all τ_i's equal and taking the $t \to \infty$ on both sides we find that while the left hand side of equation (23) is independent of t, the right hand side is the average with the equilibrium distribution (19). From here it is clear that it is not necessary to take the $t \to \infty$ limit to get the equilibrium steady state distribution — at every finite t, the stochastic correlations are already the steady-state equilibrium ones. As the steady state equilibrium distribution for the neural network system corresponds to its lowest energy state, it is clear that as $t \to \infty$, the patterns evolve into that having the lowest energy which acts as the attractor state for the dynamical system.

It is necessary to determine the global minimum of the cost function. We do this by constructing an effective cost function $\Gamma_{\mathrm{eff}}[\bar{\omega}]$ by performing a Legendre transformation which removes the dependence on J_i in favour of a dependence on the average $\bar{\omega}_i$ of the synaptic strengths:

$$\Gamma_{\mathrm{eff}}[\bar{\omega}] = -\gamma T\ln Z[J] - \int_0^\tau J_i(\tau')\bar{\omega}_i(\tau')\mathrm{d}\tau' \quad (24)$$

where

$$\bar{\omega}_i = \langle\omega_i\rangle = -\gamma T\frac{\delta\ln Z[J]}{\delta J_i} \\ J_i = -\frac{\delta\Gamma_{\mathrm{eff}}[\bar{\omega}]}{\delta\bar{\omega}_i}. \quad (25)$$

By construction, the exact effective cost function (24) is convex [10] and its global minimum gives the ground state energy of the neural network system.

Equations (24) and (25) can be combined to give

$$\Gamma_{\mathrm{eff}}[\bar{\omega}] = -\gamma T\ln\prod_{i,\mu}\int\mathcal{D}\omega_i(0)\mathrm{e}^{\frac{-E(\omega(0))}{2}}\mathcal{D}\omega_i(\tau) \\ \times\mathrm{e}^{-\frac{E(\omega(\tau))}{2}}\tilde{\mathcal{D}}\omega_i\mathrm{e}^{-\frac{1}{\gamma T}\int_0^\tau[\mathcal{L}^{\mathrm{FP}} - \frac{\delta\Gamma_{\mathrm{eff}}}{\delta\bar{\omega}_i}(\omega_i - \bar{\omega}_i)]\mathrm{d}\tau'}. \quad (26)$$

Here, just as in [7], the time interval between 0 and τ has been sliced into $N - 1$ infinitesimal parts:

$$\tilde{\mathcal{D}}\omega_i = \lim_{N\to\infty}\prod_{i=1}^{N-1}\mathcal{D}\omega_{\tau_i}, \quad (27)$$

ω_{τ_i} being the configuration of the synaptic efficacies at time τ.

As it is not possible in general to solve this exactly, we find a solution by assuming that we can make an expansion of $\Gamma_{\rm eff}[\bar{\omega}]$ in powers of a small parameter which we take as γT:

$$\Gamma_{\rm eff}[\bar{\omega}] = \sum_{n=0}^{\infty} (\gamma T)^n \Gamma^{(n)}[\bar{\omega}]. \qquad (28)$$

We discuss later how one can explicitly determine the value of the expansion parameter.

Substituting (28) into (26) and making a Taylor expansion of $\mathcal{L}^{\rm FP}$ about $\bar{\omega}$ we obtain

$$\Gamma_{\rm eff}[\bar{\omega}] = \sum_{i,\mu} \mathcal{L}^{\rm FP}(\bar{\omega}) - \gamma T \sum_{i,\mu} \mathcal{N} \ln \int \mathcal{D}\tilde{\omega}_i(0) e^{-E(\omega(0))}$$
$$\times \mathcal{D}\tilde{\omega}_i e^{-\int_0^\tau d\tau' \{\frac{1}{2}\frac{\delta^2 \mathcal{L}^{\rm FP}}{\delta \bar{\omega}_i{}^2}\tilde{\omega}_i^2 + \cdots - (\gamma T)^{1/2}\frac{\delta \Gamma^{(1)}}{\delta \bar{\omega}_i}\tilde{\omega}_i + O(\gamma T)^2\}} \quad (29)$$

where we have performed the shift:

$$(\gamma T)^{1/2}\tilde{\omega}_i = \omega_i - \bar{\omega}_i. \qquad (30)$$

By substituting the expansion (28) on the left hand side of (29), we obtain

$$\Gamma^{(0)}[\bar{\omega}] = \sum_{i,\mu} \mathcal{L}^{\rm FP}(\bar{\omega})$$
$$\Gamma^{(1)}[\bar{\omega}] = \frac{1}{2}\sum_{i,\mu} \ln \det \frac{\delta^2 \mathcal{L}^{\rm FP}}{\delta \bar{\omega}_i^2}. \qquad (31)$$

Thus the global minimum for the neural network system can be found by minimising

$$\Gamma_{\rm eff}[\bar{\omega}] = \sum_{i,\mu} \left[\mathcal{L}^{\rm FP}(\bar{\omega}) + \frac{\gamma T}{2} \ln \det \frac{\delta^2 \mathcal{L}^{\rm FP}}{\delta \bar{\omega}_i^2} + \cdots \right] \quad (32)$$

with respect to $\bar{\omega}_i$.

In the equilibrium limit the $\dot{\omega}_i$ term in (13) would not contribute to $\mathcal{L}^{\rm FP}$, so using (13) in (32) and finding the root $\bar{\omega}_{i\min}$ of the equation

$$\frac{\delta \Gamma_{\rm eff}[\bar{\omega}]}{\delta \bar{\omega}_i} = 0 \qquad (33)$$

which minimizes $\Gamma_{\rm eff}$, it is possible to avoid the problem of spurious minima for any specific choice of the cost function for a particular learning rule.

In this framework, the roots of (33) giving rise to the various local minima for the particular neural network under consideration are the source of the spurious minima for the system. Since $\Gamma_{\rm eff}[\bar{\omega}]$ is convex, it is clear that by finding the root which gives the minimal value of the effective cost function, one can immediately arrive at the global minimum of the system without getting any interference from the spurious states.

As an example, consider the cost function for a perceptron discussed in [6]:

$$E = \frac{1}{2}\sum_\mu \left(\zeta^\mu - \frac{1}{\sqrt{N}}\sum_j \omega_j \xi_j^\mu \right)^2 + \frac{\lambda}{2}\sum_j \omega_j{}^2 \quad (34)$$

where the constant λ was added to keep the connections from going to infinity. The external source J_i in our framework plays the role of the auxiliary field h_i in [6].

The Fokker-Planck Lagrangian in this case is:

$$\mathcal{L}^{\rm FP} = \frac{1}{2}\dot{\omega}_i^2 + \frac{(\gamma T)^2}{8} \left(\left(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu \xi_i^\nu \right) \omega_i \right.$$
$$\left. - \sum_{\mu\nu} \frac{1}{\sqrt{N}}\zeta^\mu \xi_i^\nu \right)^2 - \frac{\gamma T}{4} \left(\frac{1}{N}\sum_{\mu\nu}\xi_i^\nu \xi_i^\mu + \lambda \right). \quad (35)$$

Substituting $\mathcal{L}^{\rm FP}$ from (35) into (12) and using the first relation in (25) we obtain after setting $J = 0$ in the equilibrium limit,

$$\bar{\omega}_i = \frac{1}{\sqrt{N}}\xi_i^\nu \left(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu \xi_i^\nu \right)_{\nu\mu}^{-1} \zeta^\mu. \quad (36)$$

This is in agreement with the result obtained in [6].

From the fluctuation-response theorem, the response function at equilibrium is just the full connected propagator G_{ik} which is given by:

$$G_{ik}^{-1} = G_{0\,ik}^{-1} + \Sigma_{ik} \qquad (37)$$

where G_0 is the tree-level propagator and one can calculate the self-energy Σ using diagrammatic methods [5,6]. It was shown in [5,6] that the self energy is given by:

$$\Sigma = \frac{\alpha}{1+G} \qquad (38)$$

where $\alpha = p_{\max}/N$ is the storage capacity of the network. In these papers the authors calculated the storage capacity of the network at equilibrium using diagrammatic methods.

Since we have $G_{ij}^{-1} = (\gamma T)^{-1}\frac{\delta^2 \Gamma_{\rm eff}}{\delta \bar{\omega}_i \delta \bar{\omega}_j}$, we can also determine the storage capacity analytically from:

$$\alpha = 1 + (\gamma T)^{-1}\frac{\delta^2 \Gamma_{\rm eff}}{\delta \bar{\omega}^2} - \frac{\delta^2 \mathcal{L}^{\rm FP}}{\delta \bar{\omega}^2} \left(1 + \gamma T \left(\frac{\delta^2 \Gamma_{\rm eff}}{\delta \bar{\omega}^2} \right)^{-1} \right). \quad (39)$$

Since we have assumed that γT is a small parameter, it is sufficient to work up to first order in the γT expansion of $\Gamma_{\rm eff}$. The steady-state equilibrium limit $\tau \to \infty$ corresponds to the most ordered state when the patterns have settled into their attractor states. In the phase space of

synaptic strengths, it is thus a state of minimum symmetry or zero entropy. This means that in the equilibrium limit,

$$\frac{\delta}{\delta T}\left(-\gamma T \ln Z[J]\right) = \frac{\delta}{\delta T}\Big(\Gamma_{\text{eff}}[\bar{\omega}]$$
$$+ \frac{1}{\gamma T}\int_0^\tau J_i(\tau')\bar{\omega}_i(\tau')\mathrm{d}\tau'\Big) = 0. \quad (40)$$

In the thermodynamic limit $N \to \infty$, J is set to zero, and we get the result that for the attractor states:

$$\frac{\delta \Gamma_{\text{eff}}[\bar{\omega}]}{\delta T} = 0 \quad (41)$$

which shows that in this limit the effective cost function is stable with respect to changes in the noise level.

Condition (41) can be used to determine the value of the quantity γT at which the system settles into an attractor state. For the cost function (34) for example, one obtains:

$$\gamma T = \frac{1}{\sqrt{3}}\frac{\left(\frac{1}{N}\sum_{\mu\nu}\xi_i^\nu\xi_i^\mu + \lambda\right)^{1/2}}{\left[\ln\det(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu\xi_i^\nu)\right]^{1/2}} . \quad (42)$$

Applying equations (32) and (39) to the example of (34), and using equation (35), we obtain the following result for α:

$$\alpha = (1 - \gamma T)\left(1 + \frac{\gamma T}{4}\sum_i\left(\lambda + \frac{1}{N}\sum_{\mu\nu}\xi_i^\mu\xi_i^\nu\right)^2\right). \quad (43)$$

The value of γT obtained in (42) can then be substituted into (43) so that in the thermodynamic limit one obtains the result

$$\alpha = 1 + \frac{1}{\sqrt{3}}\left(\frac{\frac{1}{N}\sum_{\mu\nu}\xi_i^\nu\xi_i^\mu + \lambda}{\ln\det(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu\xi_i^\nu)}\right)^{1/2}$$
$$\times\left(\frac{1}{4}(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu\xi_i^\nu)^2 - 1\right)$$
$$-\frac{1}{12}\left(\frac{\frac{1}{N}\sum_{\mu\nu}\xi_i^\nu\xi_i^\mu + \lambda}{\ln\det(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu\xi_i^\nu)}\right)$$
$$\times\left(\lambda + \frac{1}{N}\sum_{\mu\nu}\sum_i \xi_i^\mu\xi_i^\nu\right)^2 \quad (44)$$

for the storage capacity of the network described by the cost function (34).

Although in the pseudo-inverse learning rule considered in the example above, the relation between the input and the output is linear, our construction of the effective cost function and our method for calculating the storage capacity of the network is applicable also for models with non-linear input-output relation such as:

$$\zeta^\mu = f(\sum_j \omega_j\xi_j^\mu/\sqrt{N}) \quad (45)$$

where $f(x)$ is a non-linear function of x. The procedure itself at no stage depends on any particular model and is independent of whether the input-output relation is linear or non-linear.

In this manner it is possible to calculate the storage capacity for any neural network system, in a completely analytical manner, without having to resort to diagrammatic methods, and independent of the choice of the cost function. It may be observed that the procedure we have set up in the foregoing allows the explicit calculation of the synaptic efficacies ω_{ij} from the input ξ_i^μ and the output ζ_i^μ.

3 Discussion

We have shown that a neural network system can be viewed as a non-equilibrium stochastic system of synaptic efficacies which evolve for very large values of the time into an equilibrium configuration having the lowest energy which acts as the attractor state of the network.

An effective cost function is constructed and a perturbative scheme is developed for calculating it. The global minimum of the effective cost function can be determined and gives the exact ground state energy of the system. It is shown that in the thermodynamic limit, the effective cost function is invariant under changes in the noise level. In the perturbation expansion we have constructed for Γ_{eff}, we have assumed that the expansion parameter γT is small. This can be ensured by always keeping the noise-level low so that $\gamma T \ll 1$.

In this paper we have constructed a path integral framework for the simplest case of a perceptron. It should be possible to generalize this construction for the more realistic case of many-layered neural networks and to arrive at their global minima straightaway without having to bother about the various local minima.

References

1. W.A. Little, Math. Biosci. **19**, 101 (1974).
2. E. Gardner, J. Phys. A **21**, 257 (1988).
3. E. Gardner, B. Derrida, J. Phys. A **21**, 271 (1988).
4. B. Muller, J. Reinhardt, *Neural Networks: An Introduction* (Springer-Verlag, 1990).
5. J.A. Hertz, G.I. Thorbergsson, Phys. Scr. **T25**, 149 (1989).
6. J.A. Hertz, A. Krogh, G.I. Thorbergsson, J. Phys. A **22**, 2133 (1989).
7. H. Sompolinsky, A. Zippelius, Phys. Rev. B **25**, 6860 (1982); A.Crisanti, H. Sompolinsky, Phys. Rev. A **36**, 4922 (1987); R. Kree, A. Zippelius, Phys. Rev. A **36**, 4421 (1987); H. Sompolinsky, N. Tishby, H.S. Seung, Phys. Rev. Lett. **65**, 1683 (1990).
8. E. Gozzi, Phys. Rev. D **28**, 1922 (1983).
9. E. Gozzi, Phys. Lett. B **130**, 183 (1983).
10. J. Iliopoulos, C. Itzykson, A. Martin, Rev. Mod. Phys. **47**, 165 (1975).