

## Thouless-Anderson-Palmer equation and self-consistent signal-to-noise analysis for the Hopfield model with three-body interaction

Akihisa Ichiki\* and Masatoshi Shiino

*Department of Applied Physics, Faculty of Science, Tokyo Institute of Technology, 2-12-1 Ohokayama Meguro-ku Tokyo, Japan*

(Received 2 March 2006; published 25 July 2006)

The self-consistent signal-to-noise analysis (SCSNA) is an alternative to the replica method for deriving the set of order parameter equations for associative memory neural network models and is closely related with the Thouless-Anderson-Palmer equation (TAP) approach. In the recent paper by Shiino and Yamana the Onsager reaction term of the TAP equation has been found to be obtained from the SCSNA for Hopfield neural networks with two-body interaction. We study the TAP equation for an associative memory stochastic analog neural network with three-body interaction to investigate the structure of the Onsager reaction term, in connection with the term proportional to the output characteristic to the SCSNA. We report on the SCSNA framework for analog networks with three-body interactions as well as provide a recipe based on the cavity concept that involves two cavities and the hybrid use of the SCSNA to obtain the TAP equation.

DOI: [10.1103/PhysRevE.74.017103](https://doi.org/10.1103/PhysRevE.74.017103)

PACS number(s): 84.35.+i, 75.10.Nr, 87.18.Sn

The replica method [1] for random spin systems has been successfully employed in neural network models of associative memory [2,3]. However, the replica calculations require the concept of free energy. On the other hand an alternative approach to obtain the order parameter equations of the self-consistent signal-to-noise analysis (SCSNA) [4,5] for deterministic analog neural networks is free from the energy concept and thus applicable to networks with asymmetric connections. The SCSNA was shown to be closely related to the Thouless-Anderson-Palmer (TAP) equation approach [2,6–9] through the cavity concept in the case where systems have an energy function [10].

An advantage of dealing with the TAP equation of neural networks is that equilibrium behaviors of a stochastic neural network can be studied by investigating the corresponding TAP equation that is viewed as a deterministic analog network [5]. The set of order parameter equations of the original stochastic network is obtained by applying the SCSNA to the TAP equation.

TAP equations, which have recently been attracting much attention from the viewpoint of applying statistical mechanics to information theoretic engineering problems [11,12], are usually obtained by means of the Plefka method [13]. A more systematic method of deriving TAP equations incorporates the cavity concept [2,8] to elucidate the structure as well as the meaning of TAP equations. Noting that the SCSNA and the TAP equation approach share a common idea of the cavity concept, Shiino and Yamana have found that the Onsager reaction term characteristic to a TAP equation can be recovered by the SCSNA for stochastic analog networks with two-body interactions [10]. The aim of the present paper is studying the SCSNA framework and the TAP equation of stochastic analog networks with multibody interaction to elucidate the structure of the Onsager reaction term in connection with the term proportional to the output characteristic to the SCSNA.

The Ising spin Hopfield models with  $p$ -body connections

[14] which are analogous to  $p$ -spin random spin glass models [15] were studied to explore the statistical behavior of their retrieval properties. It is well known that the storage capacity of the network is proportional to  $N^{p-1}$  [14] where  $N$  represents the number of neurons. In the present paper we report on a method based on the cavity approach [2] to derive the TAP equations for the  $p=3$  Hopfield model with simultaneous use of the SCSNA for this model. Since we deal with analog neurons, or soft spins in the present paper, we can choose transfer functions of arbitrary shapes. The TAP equations for analog networks with multibody interactions have not been reported.

We deal with a stochastic analog neural network of the form

$$\frac{dx_i}{dt} = -\frac{d\phi(x_i)}{dx_i} + \sum_{j < k (\neq i)} J_{ijk} x_j x_k + \eta_i(t), \quad (1)$$

where  $x_i$  represents a state of an analog neuron or a soft spin at site  $i$ ,  $\phi(x_i)$  the potential,  $\eta_i$  the Langevin white noise obeying  $\langle \eta_i(t) \eta_j(t') \rangle = \frac{2}{\beta} \delta(t-t') \delta_{ij}$  for  $i=1 \cdots \tilde{N}$ ,  $\beta$  the intensity of externally driven Langevin noise and the synaptic coupling  $J_{ijk}$  is assumed to be given by the Hebb learning rule extended to the  $p=3$  Hopfield model [14]:

$$J_{ijk} = \frac{1}{\tilde{N}^2} \sum_{\mu=1}^{\tilde{p}} \xi_i^\mu \xi_j^\mu \xi_k^\mu \quad (2)$$

for  $i \neq j \neq k$  and otherwise  $J_{ijk}=0$  with  $\xi_i^\mu = \pm 1$  representing  $\tilde{p}(=\alpha\tilde{N}^2)$  random memory patterns.

For the  $p=3$  Hopfield model, the local field at site  $i$  is defined as  $h_i = \sum_{j < k (\neq i)} J_{ijk} x_j x_k$  and we have to calculate the second moments of soft spins  $\langle x_j x_k \rangle$  to obtain the TAP equation which is given by expressing the thermal averages of the local fields in terms of those of soft spins. For this reason the standard cavity method applied to networks with  $p=2$ , where it suffices to take only one cavity into account, is ineffective

\*Electronic address: [aichiki@mikan.ap.titech.ac.jp](mailto:aichiki@mikan.ap.titech.ac.jp)

in the present case. To study the TAP equation for the system (1) we have to introduce two cavities. Hereafter we will refer to the sites  $i=0, -1$  as the cavity sites.

The Fokker-Planck equation for the probability density of soft spins corresponding to the set of Langevin equations (1) with  $\tilde{N}=N+2$  has the equilibrium probability density  $P_{\text{eq}}(x_{-1}, x_0, x_1, \dots, x_N) = \frac{1}{Z_{N+2}} \exp(-\beta H^{(N+2)})$ , where  $Z_{N+2}$  is the partition function of the system including two cavities and  $H^{(N+2)}$  is its Hamiltonian. Noting that the Hamiltonian of  $N$ -body system  $\{x_1, \dots, x_N\}$  is given as follows:

$$H^{(N)} = \sum_{i=1}^N \phi(x_i) - \sum_{i < j < k=1}^N J_{ijk} x_i x_j x_k, \quad (3)$$

the total Hamiltonian  $H^{(N+2)}$  of the system which includes the effects of cavity neurons is expressed as follows:

$$H^{(N+2)} = H^{(N)} + \phi(x_0) + \phi(x_{-1}) - \tilde{h}_0 x_0 - \tilde{h}_{-1} x_{-1} - \hat{h}_{-10} x_0 x_{-1}, \quad (4)$$

where  $\tilde{h}_0$ ,  $\tilde{h}_{-1}$  and  $\hat{h}_{-10}$  are, respectively, defined

$$\tilde{h}_0 = \frac{1}{2} \frac{1}{(N+2)^2} \sum_{\mu} \sum_{\substack{i,k=1 \\ j \neq k}}^N \xi_0^{\mu} \xi_j^{\mu} \xi_k^{\mu} x_j x_k, \quad (5)$$

$$\tilde{h}_{-1} = \frac{1}{2} \frac{1}{(N+2)^2} \sum_{\mu} \sum_{\substack{i,k=1 \\ j \neq k}}^N \xi_{-1}^{\mu} \xi_j^{\mu} \xi_k^{\mu} x_j x_k, \quad (6)$$

$$\hat{h}_{-10} = \frac{1}{(N+2)^2} \sum_{\mu} \sum_{k=1}^N \xi_{-1}^{\mu} \xi_0^{\mu} \xi_k^{\mu} x_k. \quad (7)$$

Notice that  $\tilde{h}_0$  and  $\tilde{h}_{-1}$  are the local fields induced by the  $N$  neurons at cavity sites  $i=0, -1$ , respectively, and  $\hat{h}_{-10}$  is the local field of another type affecting the pair of cavities. Then using the marginal probability density and noting that the marginal probability density of the local fields of the  $N$ -body system is expected to be a three-dimensional Gaussian distribution since  $\tilde{h}_0$ ,  $\tilde{h}_{-1}$ , and  $\hat{h}_{-10}$  are the sums of independent random variables, one can evaluate the  $(N+2)$ -body average of the cavity soft spin  $x_0$  as follows:

$$\langle x_0 \rangle_{N+2} = F(\langle \tilde{h}_0 \rangle_N) + (\beta \langle \hat{h}_{-10} \rangle_N + \beta^2 \sigma_{-10}^2) [G(\langle \tilde{h}_0 \rangle_N) - F^2(\langle \tilde{h}_0 \rangle_N)] F(\langle \tilde{h}_{-1} \rangle_N) + O(1/N), \quad (8)$$

where  $F$  is the transfer function which is defined

$$F(y) \equiv \int dx x \exp \left[ -\beta \phi(x) + \beta y x + \frac{\beta^2 \sigma^2 x^2}{2} \right] \times \left\{ \int dx \exp \left[ -\beta \phi(x) + \beta y x + \frac{\beta^2 \sigma^2 x^2}{2} \right] \right\}^{-1} \quad (9)$$

and

$$G(y) \equiv \int dx x^2 \exp \left[ -\beta \phi(x) + \beta y x + \frac{\beta^2 \sigma^2 x^2}{2} \right] \times \left\{ \int dx \exp \left[ -\beta \phi(x) + \beta y x + \frac{\beta^2 \sigma^2 x^2}{2} \right] \right\}^{-1}, \quad (10)$$

$\sigma^2$  the variance of the local field  $\tilde{h}_i$ ,  $\sigma_{-10}^2$  the covariance of  $\tilde{h}_0$  and  $\tilde{h}_{-1}$ , and  $\langle \cdot \rangle_N$  the  $N$ -body average. Since  $\langle \tilde{h}_{-10} \rangle_N = O(1/\sqrt{N})$ ,  $\sigma_{-10}^2 = O(1/\sqrt{N})$ ,

$$\langle \tilde{h}_0 \rangle_{N+2} = \langle \tilde{h}_0 \rangle_N + \beta \sigma^2 \langle x_0 \rangle_{N+2} + O(1/\sqrt{N}), \quad (11)$$

and the true local field  $h_0$  is given by  $h_0 = \tilde{h}_0 + \hat{h}_{-10} x_{-1}$ , we have

$$\langle x_0 \rangle_{N+2} = F(\langle h_0 \rangle_{N+2} - \beta \sigma^2 \langle x_0 \rangle_{N+2}). \quad (12)$$

This equation coincides with the one obtained using the standard cavity method for the Hopfield model where one cavity is taken.

However, it does not suffice to obtain the TAP equation in the present case, because  $\langle h_0 \rangle_{N+2}$  is simply given in terms of the second moments of soft spins  $\langle x_j x_k \rangle_{N+2}$ . In order to find the expression of the local field  $\langle h_i \rangle_{N+2}$  in terms of the thermal averages of soft spins, it will be necessary to employ a two-cavity method and we obtain

$$\langle x_0 x_{-1} \rangle_{N+2} = \langle x_0 \rangle_{N+2} \langle x_{-1} \rangle_{N+2} + (\beta \langle \hat{h}_{-10} \rangle_{N+2} + \beta^2 \sigma_{-10}^2) \times (\langle x_0^2 \rangle_{N+2} - \langle x_0 \rangle_{N+2}^2) (\langle x_{-1}^2 \rangle_{N+2} - \langle x_{-1} \rangle_{N+2}^2) - \langle x_{-1} \rangle_{N+2}^2 + O(1/N). \quad (13)$$

Thus the thermal average of the local field  $\langle h_i \rangle_{N+2}$  is given as follows:

$$\langle h_i \rangle_{N+2} = \sum_{j < k (\neq i)} J_{ijk} \langle x_j \rangle_{N+2} \langle x_k \rangle_{N+2} + \frac{\alpha}{2\beta} \langle x_i \rangle_{N+2} U^2 + O(1/\sqrt{N}), \quad (14)$$

where  $U$  is the average of susceptibilities

$$U = \frac{\beta}{N+2} \sum_j (\langle x_j^2 \rangle_{N+2} - \langle x_j \rangle_{N+2}^2). \quad (15)$$

Notice that  $\langle \hat{h}_{ij} \rangle_N$  affects the  $(N+2)$ -body average of the local field  $\langle h_i \rangle_{N+2}$  via the summation with respect to the site index, on the other hand  $\sigma_{-10}^2$  does not affect the local field. Equation (14) together with Eq. (12) yields the pre-TAP equations [10] for a  $p=3$  Hopfield model;

$$\langle x_i \rangle_N = F \left( \sum_{j < k (\neq i)} J_{ijk} \langle x_j \rangle_N \langle x_k \rangle_N - \Gamma_{\text{TAP}} \langle x_i \rangle_N \right) \quad (i = 1, \dots, N) \quad (16)$$

with

$$\Gamma_{\text{TAP}} = \beta\sigma^2 - \frac{\alpha U^2}{2\beta}, \quad (17)$$

where  $\langle x_i \rangle_N$  is used instead of  $\langle x_i \rangle_{N+2}$ .  $-\Gamma_{\text{TAP}} \langle x_i \rangle_N$  is the Onsager reaction term of the TAP equation. However,  $\Gamma_{\text{TAP}}$ , especially,  $\sigma^2$  in Eq. (9) has still to be determined to obtain the TAP equation.

Equation (16) defines a deterministic analog network of associative memory with the transfer function  $F$  corresponding to the original stochastic analog neural network (1). To obtain  $\sigma^2$  together with the order parameter equations we then apply the SCSNA to Eq. (16). The SCSNA is a self-consistent method for properly renormalizing the so-called noise part due to interference of noncondensed patterns in the local field of a neuron  $h_i^{\text{det}} \equiv \sum_{j < k (\neq i)} J_{ijk} \langle x_j \rangle_N \langle x_k \rangle_N$  in Eq. (16), which can be rewritten

$$h_i^{\text{det}} = \frac{1}{2} \sum_{\mu=1}^{\alpha N^2} \xi_i^\mu \left[ (m^\mu)^2 - \frac{1}{N} \sum_j \frac{\langle x_j \rangle_N^2}{N} \right] - \frac{\langle x_i \rangle_N}{N} \sum_{\mu=1}^{\alpha N^2} m^\mu + \frac{\langle x_i \rangle_N^2}{N^2} \sum_{\mu=1}^{\alpha N^2} \xi_i^\mu, \quad (18)$$

where  $m^\mu$  is the overlap of  $\mu^{\text{th}}$  pattern defined as  $m^\mu \equiv \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \langle x_i \rangle_N$  for  $\mu=1, \dots, \alpha N^2$ . To extract the pure noise obeying a Gaussian distribution, we decompose the local field  $h_i^{\text{det}}$  assuming that the first pattern is condensed and the others are noncondensed, i.e.,  $m^1 = O(1)$  and  $m^\mu = O(1/\sqrt{N})$  ( $\mu \geq 2$ ) as follows [4,5]:

$$h_i^{\text{det}} = \frac{1}{2} \xi_i^1 (m^1)^2 + \frac{1}{2} \xi_i^\mu \left[ (m^\mu)^2 - \frac{\tilde{C}}{N} \right] - \frac{\langle x_i \rangle_N}{N} m^\mu + \frac{\langle x_i \rangle_N^2}{N^2} \xi_i^\mu + z_{i\mu} + \Gamma_{\text{SCSNA}} \langle x_i \rangle_N, \quad (19)$$

where

$$\tilde{C} \equiv \frac{1}{N} \sum_{i=1}^N \langle x_i \rangle_N^2, \quad (20)$$

$$\Gamma_{\text{SCSNA}} \equiv \gamma - \tilde{\gamma}, \quad (21)$$

$$\tilde{\gamma} \equiv \frac{1}{N} \sum_{\mu \geq 2} m^\mu, \quad (22)$$

$$z_{i\mu} + \gamma \langle x_i \rangle_N \equiv \frac{1}{2} \sum_{\nu \neq 1, \mu} \xi_i^\nu \left[ (m^\nu)^2 - \frac{\tilde{C}}{N} \right], \quad (23)$$

and  $z_{i\mu}$  is assumed to be a Gaussian random variable with mean zero and the variance is to be evaluated self-consistently.

Substituting the expression of the local field Eq. (19) into the pre-TAP equation (16) and comparing with  $\langle x_i \rangle_N = F(\langle \tilde{h}_i \rangle_{N-2})$ , which holds by observing Eq. (8) for large  $N$ , it follows that  $\Gamma_{\text{SCSNA}} = \Gamma_{\text{TAP}}$  [10], since  $\langle \tilde{h}_i \rangle_{N-2} - \frac{1}{2} \xi_i^1 (m^1)^2$  obeys a Gaussian random variable. Then we have

$$\langle x_i \rangle_N = F \left( \frac{1}{2} \xi_i^1 (m^1)^2 + \frac{1}{2} \xi_i^\mu \left[ (m^\mu)^2 - \frac{\tilde{C}}{N} \right] + z_{i\mu} \right) \quad (24)$$

from which it follows [5]:

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu F \left( \frac{1}{2} \xi_i^1 (m^1)^2 + z_{i\mu} \right) + \frac{1}{2} \left[ (m^\mu)^2 - \frac{\tilde{C}}{N} \right] \times \frac{1}{N} \sum_{i=1}^N F' \left( \frac{1}{2} \xi_i^1 (m^1)^2 + z_{i\mu} \right) + O(N^{-5/2}), \quad (25)$$

where  $F'$  denotes the derivative of the transfer function  $F$ .

It should be noted that in the present case it is necessary to obtain  $m^\mu (= O(1/\sqrt{N}))$  up to  $O(1/N)$  unlike the case of  $p=2$  where up to  $O(1/\sqrt{N})$  of  $m^\mu$  suffices. We solve this equation for  $m^\mu$  perturbatively by putting  $m^\mu = m_{1/2}^\mu + m_1^\mu + O(N^{-3/2})$ , where  $m_k^\mu$  represents the part of  $O(N^{-k})$ . Substituting the solution of Eq. (25) into the definitions of  $\Gamma_{\text{SCSNA}}$  and  $z_{i\mu}$ , we find  $\tilde{\gamma} = 0$  in the limit of large  $N$ . Furthermore,  $\Gamma_{\text{SCSNA}}$  which determines the form of the transfer function  $F$  by the relation

$$\Gamma_{\text{SCSNA}} = \Gamma_{\text{TAP}} = \beta\sigma^2 - \frac{\alpha U^2}{2\beta} \quad (26)$$

is evaluated using the self-averaging property

$$\Gamma_{\text{SCSNA}} = \alpha U \left\langle F^2 \left( \frac{1}{2} \xi^1 (m^1)^2 + z \right) \right\rangle_{\xi^1, z}, \quad (27)$$

where  $U$  is given as follows:

$$U = \left\langle F' \left( \frac{1}{2} \xi^1 (m^1)^2 + z \right) \right\rangle_{\xi^1, z} \quad (28)$$

and  $\langle \cdot \rangle_{\xi^1, z}$  represents the average over the random pattern  $\xi^1 = \pm 1$  and the Gaussian random variable  $z$ . Notice that  $U$  given by this equation corresponds to that defined by Eq. (15), since differentiating Eq. (9) with respect to  $y$  yields the susceptibility and the self-averaging property holds. Similarly the variance of the Gaussian noise  $z$  can be calculated as follows:

$$\sigma_z^2 = \frac{\alpha}{2} \left\langle F^2 \left( \frac{1}{2} \xi^1 (m^1)^2 + z \right) \right\rangle_{\xi^1, z}^2. \quad (29)$$

Furthermore the overlap of the condensed pattern  $m^1$  is given as follows:

$$m^1 = \left\langle \xi^1 F \left( \frac{1}{2} \xi^1 (m^1)^2 + z \right) \right\rangle_{\xi^1, z}. \quad (30)$$

Equations (27)–(30) constitute the SCSNA framework in terms of  $m^1$ ,  $U$ ,  $\sigma^2$ , and  $\sigma_z^2$  that yields the order parameter equations for determining the storage capacity of the analog network (16) and hence of the stochastic analog neural network (1). We note that they coincide with those obtained by the replica calculation when the replica symmetry is assumed to hold. Furthermore, in the special case where  $\exp -\beta\phi(x) / \int \exp -\beta\phi(x) dx = \frac{1}{2} (\delta(x+1) + \delta(x-1))$ , the result of the SCSNA recovers that of the Ising spin network studied

in [14]. Then, from Eqs. (16) and (17) we finally obtain the TAP equation

$$\langle x_i \rangle_N = F \left( \sum_{j < k (\neq i)} J_{ijk} \langle x_j \rangle_N \langle x_k \rangle_N - \Gamma_{\text{SCSNA}} \langle x_i \rangle_N \right) \quad (31)$$

with  $\Gamma_{\text{SCSNA}}$  and the form of the transfer function  $F$  self-consistently determined by the above Eqs. (26)–(30). Noting that  $\Gamma_{\text{SCSNA}}$  can be expressed in terms of an Edwards-Anderson order parameter  $q$ , the TAP equation (31) can be further rewritten as  $\langle x_i \rangle_N = F(\sum_{j < k (\neq i)} J_{ijk} \langle x_j \rangle_N \langle x_k \rangle_N - \alpha \beta q (\hat{q} - q) \langle x_i \rangle_N)$ , where the Edward-Anderson order parameter  $q$  and  $\hat{q}$  are defined, respectively, as  $q = \frac{1}{N} \sum_i \langle x_i \rangle_N^2$  and  $\hat{q} = \frac{1}{N} \sum_i \langle x_i^2 \rangle_N$  which are easily computed using Eqs. (9) and (10).

In conclusion, taking advantage of the close relationship between the TAP equation and the SCSNA approaches, we have studied the SCSNA framework and the TAP equation for the stochastic analog network with three-body interaction based on the Hebb learning rule. Characteristic to the TAP equation for the case of such an interaction is that the Onsager reaction term consists of the two terms, as is shown in Eq. (17): the one arising from the variance of the Gaussian distribution for the local fields that determines the shape of

the transfer function  $F$  and the other one due to the two body correlation of soft spins the computation of which requires taking two cavities. Without resorting to the Hamiltonian based on overlap evaluation by adding a memory pattern to the network, which is usually taken for the Hopfield model [2,10], the coefficient of the Onsager reaction term has been given within the framework of the SCSNA, although applying the conventional recipe recovers the result. Renormalization of the noise part of the local field in the SCSNA scheme corresponds to the overlap evaluation in the pattern-adding approach of the cavity method. The framework of the SCSNA described in its application to the (pre-)TAP equation of our stochastic network implies that the set of order parameter equations obtained there still formally makes sense in the case of general deterministic analog networks having three-body interaction and transfer function  $F$  of arbitrary shape irrespective of whether it is monotonic or not. Details of the analysis including the phase diagram showing the behavior of the storage capacity will be published elsewhere.

This work was supported by a 21st Century COE Program at TokyoTech “Nanometer-Scale Quantum Physics” by the Ministry of Education, Culture, Sports, Science and Technology.

- 
- [1] D. Sherrington and S. Kirkpatrick, *Phys. Rev. Lett.* **35**, 1792 (1975).  
 [2] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).  
 [3] D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. Lett.* **55**, 1530 (1985).  
 [4] M. Shiino and T. Fukai, *J. Phys. A* **25**, L375 (1992).  
 [5] M. Shiino and T. Fukai, *Phys. Rev. E* **48**, 867 (1993).  
 [6] D. J. Thouless, P. W. Anderson, and R. G. Palmer, *Philos. Mag.* **35**, 593 (1977).  
 [7] T. Morita and T. Horiguchi, *Solid State Commun.* **19**, 833 (1976).  
 [8] M. Shamir and H. Sompolinsky, *Phys. Rev. E* **61**, 1839 (2000).  
 [9] M. Mézard, *J. Phys. A* **22**, 2181 (1989).  
 [10] M. Shiino and M. Yamana, *Phys. Rev. E* **69**, 011904 (2004).  
 [11] *Learning in Graphical Models*, edited by M. Jordan (MIT Press, Cambridge, 1999).  
 [12] *Advanced Mean Field Methods, Theory and Practice*, edited by M. Opper and D. Saad (MIT Press, Cambridge, 2001).  
 [13] T. Plefka, *J. Phys. A* **15**, 1971 (1982).  
 [14] L. F. Abbott and Y. Arian, *Phys. Rev. A* **36**, 5091 (1987).  
 [15] H. Rieger, *Phys. Rev. B* **46**, 14655 (1992).