Thouless-Anderson-Palmer Equation and Self-Consistent Signal-to-Noise Analysis for Hopfield Model with Multibody Interaction

Akihisa Ichiki and Masatoshi Shiino

Department of Applied Physics, Faculty of Science, Tokyo Institute of Technology, 2-12-2 Ohokayama Meguro-ku Tokyo, Japan
aichiki@mikan.ap.titech.ac.jp

Abstract. 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 networks 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 2-body interaction. We study the TAP equation for an associative memory stochastic analog neural network with 3-body interaction to investigate the structure of the Onsager reaction term, in connection with the output proportional term characteristic to the SCSNA. We report the SCSNA framework for analog networks with 3-body interactions as well as a novel recipe based on the cavity concept that involves two cavities and the hybrid use of the SCSNA to obtain the TAP equation.

1 Self-Consistent Signal-to-Noise Analysis

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 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 2-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 a novel 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 the equilibrium state of a deterministic analog network defined as $x_i = F(h_i)$, where $x_i$ is a state of an analog neuron at site $i$, $h_i \equiv \sum_{j<k} J_{ijk} x_j x_k$ is the local field of the site $i$, $F$ is a transfer function of an arbitrary shape and the synaptic coupling $J_{ijk}$ is assumed to be given by the Hebb learning rule extended to $p = 3$ Hopfield model [14] as $J_{ijk} = \frac{1}{N^3} \sum_{\mu=1}^{\alpha N^2} \xi_{i}^{\mu} \xi_{j}^{\mu} \xi_{k}^{\mu}$ for $i \neq j \neq k$ and otherwise $J_{ijk} = 0$ with $\xi_{i}^{\mu} = \pm 1$ representing $\tilde{p}(= \alpha N^2)$ random memory patterns.

The SCSNA is a self-consistent method for properly renormalizing the so-called noise part due to interference of non-condensed patterns in the local field of a neuron $h_i \equiv \sum_{j<k(\neq i)} J_{ijk} x_j x_k$ in a deterministic analog network, which can be rewritten as

$$h_i = \frac{1}{2} \sum_{\mu=1}^{\alpha N^2} \xi_{i}^{\mu} \left( m_{\mu}^2 - \frac{1}{N} \sum_{j} x_j^2 \right) - \frac{x_i}{N} \sum_{\mu=1}^{\alpha N^2} m_{\mu}^2 + \frac{x_i^2}{N^2} \sum_{\mu=1}^{\alpha N^2} \xi_{i}^{\mu}, \quad (1)$$

where $m_{\mu}$ is the overlap of $\mu^{th}$ pattern defined as $m_{\mu} \equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \xi_{i}^{\mu} x_{i}$ ($\mu = 1, \ldots, \alpha N^2$). To extract the pure noise obeying a Gaussian distribution, we decompose the local field $h_i$ assuming that the first pattern is condensed and the others are non-condensed, i.e., $m^1 = O(1)$ and $m_{\mu} = O(1/\sqrt{N})$ ($\mu \geq 2$) as [4,5]

$$h_i = \frac{1}{2} \xi_{i}^{1} \left( m^1 \right)^2 + \frac{1}{2} \xi_{i}^{\mu} \left( m_{\mu}^2 - \frac{C}{N} \right) - \frac{x_i}{N} m_{\mu}^2 + \frac{x_i^2}{N^2} \xi_{i}^{\mu} + z_{\mu} + T_{\text{SCSNA}} x_i, \quad (2)$$
where $\tilde{C} = \frac{1}{N} \sum_{i=1}^{N} x_i^2$, $\Gamma_{SCSNA} = \gamma - \tilde{\gamma}$, $\tilde{\gamma} = \frac{1}{N} \sum_{\mu \geq 2} m^\mu$, $z_{i\mu} + \gamma x_i \equiv \frac{1}{2} \sum_{\nu \neq 1, \mu} \xi_i^\mu \{(m^\nu)^2 - \tilde{C}/N\}$, 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.(2) into $x_i = F(h_i)$ and solving this equation with respect to $x_i$, we have

$$x_i = \tilde{F} \left( \frac{1}{2} \xi_i^1 (m^1)^2 + \frac{1}{2} \xi_i^\mu \left( (m^\mu)^2 - \frac{\tilde{C}}{N} \right) + z_{i\mu} \right),$$

from which it follows [5]

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

where $\tilde{F}(y) \equiv F(y + \Gamma_{SCSNA} \tilde{F}(y))$ and $\tilde{F}'$ denotes the derivative of the transfer function $\tilde{F}$. Noting that in the present case it is necessary to obtain $m^\mu$ 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 $m^\mu$ perturbedly by putting $m^\mu = m^\mu_{1/2} + m^\mu_1 + O(N^{-3/2})$, where $m^\mu_{1/2}$ represents the part of $O(N^{-k})$. Then we find $m^\mu_{1/2} = \tilde{M}_\mu$, $m^\mu_1 = \frac{U_{m}}{2} \xi_i^1 (m^1)^2 - \frac{\tilde{C}}{N}$, where $\tilde{M}_\mu \equiv \frac{1}{N} \sum_{i} \xi_i^\mu \tilde{F} \left( \frac{1}{2} \xi_i^1 (m^1)^2 + z_{i\mu} \right)$ and $U_{m} \equiv \frac{1}{N} \sum_{i} \tilde{F} \left( \frac{1}{2} \xi_i^1 (m^1)^2 + z_{i\mu} \right)$. Substituting these equations into the equations defining $\Gamma_{SCSNA}$, $\tilde{\gamma}$ and $z_{i\mu}$, we find $\tilde{\gamma} = 0$ in the limit of large $N$, $\Gamma_{SCSNA} = \lim_{N \to \infty} \frac{1}{4} \sum_{\mu} \xi_i^\mu \{ (M_\mu)^3 - \frac{1}{4} \tilde{M}_\mu \tilde{C} \} U_{m}$, $z_{i\mu} = \frac{1}{\sqrt{N}} \sum_{\nu \neq 1, \mu} \sum_{j \neq k(\neq i)} \xi_j^\mu \xi_k^\nu \tilde{F} \left( \frac{1}{2} \xi_k^1 (m^1)^2 + z_{j\nu} \right) \tilde{F} \left( \frac{1}{2} \xi_k^1 (m^1)^2 + z_{k\nu} \right)$. Furthermore using the self-averaging property, we note that $U_{m}$ is independent of the pattern index $\mu$ to have

$$U_{m} = U = \left\langle \tilde{F}' \left( \frac{1}{2} \xi_1^1 (m^1)^2 + z \right) \right\rangle_{\xi_1^1, z},$$

where $\langle \rangle_{\xi_1^1, z}$ represents the average over the random pattern $\xi_1^1 = \{ \pm 1 \}$ and the Gaussian random variable $z$.

$\Gamma_{SCSNA}$ which determines the form of the transfer function $\tilde{F}$ is evaluated using the self-averaging property as

$$\Gamma_{SCSNA} = \alpha U \left\langle \tilde{F}^2 \left( \frac{1}{2} \xi_1^1 (m^1)^2 + z \right) \right\rangle_{\xi_1^1, z}.$$  

Similarly the variance of the Gaussian noise $z$ can be calculated using the expression of $z$ as

$$\sigma_z^2 = \alpha \frac{2}{\alpha} \left\langle \tilde{F}^2 \left( \frac{1}{2} \xi_1^1 (m^1)^2 + z \right) \right\rangle_{\xi_1^1, z}^2.$$
Furthermore the overlap of the condensed pattern $m^1$ is given as

$$m^1 = \left\langle \xi^1 \hat{F} \left( \frac{1}{2} \xi^1 (m^1)^2 + z \right) \right\rangle_{\xi^1, z}. \quad (8)$$

Equations (5), (6), (7) and (8) constitute the SCSNA framework that yields the order parameter equations for determining the storage capacity of the network $x_i = F(h_i)$.

2 Cavity Method

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 wellknown that the storage capacity of the network is proportional to $N^{p-1}$ [14] where $N$ represents the number of neurons. In this section we will see a novel 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.

Although we have dealt with the deterministic network in the previous section, we deal with a stochastic analog neural network in this section:

$$\frac{dx_i}{dt} = -\frac{d\phi(x_i)}{dx_i} + \sum_{j<k \neq i} J_{ijk} x_j x_k + \eta_i(t), \quad (9)$$

where $\langle \eta_i(t)\eta_j(t') \rangle = \frac{2}{\beta} \delta(t - t') \delta_{ij}$ ($i = 1 \cdots \tilde{N}$), $x_i$ represents a state of an analog neuron at site $i$, $\phi(x_i)$ the potential, $\beta$ the intensity of externally driven Langevin noise, $\tilde{N}$ the number of neurons and the synaptic coupling $J_{ijk}$ is again assumed to be given by the Hebb learning rule.

For the $p = 2$ Hopfield model with analog neurons, the thermal average of a state of a neuron, in general, is given by the set of the equations $\langle x_i \rangle = F(\langle h_i \rangle - \Gamma_{\text{Ons}} \langle x_i \rangle)$, where $h_i$ is a local field at site $i$ generated by other neurons, $F$ a transfer function and $\Gamma_{\text{Ons}} \langle x_i \rangle$ is called the Onsager reaction term. Since the local field at site $i$ is defined as the linear combination of $x_i$’s, i.e., $h_i = \sum_{j \neq i} J_{ij} x_j$ for $p = 2$ Hopfield model, $\langle x_i \rangle = F(\langle h_i \rangle - \Gamma_{\text{Ons}} \langle x_i \rangle)$ with the definition of the local field yields the TAP equation which gives the thermal average of a state of a neuron.

For the $p = 3$ Hopfield model, however, 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 states of neurons $\langle x_j x_k \rangle$ to obtain the TAP equations which are given by expressing the thermal averages of the local fields in terms of those of states of neurons $\langle x_i \rangle$’s. For this reason the cavity method applied to networks with 2-body interaction, where only one cavity is taken does not suffice in the present case. To study the TAP equation for the system (9) 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 states of neurons corresponding to the set of Langevin equations (9) with $N = N + 2$ has the equilibrium probability density $P_{eq}(x_{-1}, x_0, x_1, \cdots, x_N) = \frac{1}{Z_{N+2}} \exp(-\beta H^{(N+2)})$, 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 \{ $x_1, \cdots, x_N$ \} is given as $H^{(N)} = \sum_{i=1}^{N} \phi(x_i) - \sum_{i<j<k=1}^{N} J_{ijk} x_i x_j x_k$, the total Hamiltonian $H^{(N+2)}$ of the system which includes the effects of cavity neurons is expressed as $H^{(N+2)} = H^{(N)} + \phi(x_0) + \phi(x_{-1}) - h_0 x_0 - h_{-1} x_{-1} - h_{-10} x_0 x_{-1}$, where $h_0$, $h_{-1}$ and $h_{-10}$ are respectively defined as $h_0 = \frac{1}{2(N+2)^2} \sum_{\mu} \sum_{j \neq k=1}^{N} \xi_{\mu}^j \xi_{k}^j x_j x_k$ and $h_{-10} = \frac{1}{(N+2)^2} \sum_{\mu} \sum_{k=1}^{N} \xi_{\mu}^0 \xi_{k}^0 x_k$. Notice that $h_0$ and $h_{-1}$ are the local fields induced by the $N$ neurons at cavity sites $i = 0, -1$ respectively and $h_{-10}$ is the local field of another type affecting the pair of cavities. Then the marginal probability density $P_{N+2}(x_0, \tilde{h}_0; x_{-1}, \tilde{h}_{-1}; \tilde{h}_{-10})$ of the $(N + 2)$-body equilibrium probability density $P_{eq}(x_{-1}, x_0, x_1, \cdots, x_N)$ reads $P_{N+2}(x_0, \tilde{h}_0; x_{-1}, \tilde{h}_{-1}; \tilde{h}_{-10}) = C \exp[-\beta(\phi(x_0) + \phi(x_{-1}) - h_0 x_0 - h_{-1} x_{-1} - h_{-10} x_0 x_{-1})] P_N(\tilde{h}_0, \tilde{h}_{-1}, \tilde{h}_{-10})$, where $C$ is the normalization constant and $P_N(\tilde{h}_0, \tilde{h}_{-1}, \tilde{h}_{-10})$ is the marginal probability density of the $N$-body equilibrium density:

$$P_N(\tilde{h}_0, \tilde{h}_{-1}, \tilde{h}_{-10}) = \frac{1}{Z_N} \int \prod_{i=1}^{N} dx_i e^{-\beta H^{(N)}} \delta \left( \tilde{h}_0 - \frac{1}{2} \sum_{j,k=1}^{N} J_{0jk} x_j x_k \right) \times \delta \left( \tilde{h}_{-1} - \frac{1}{2} \sum_{j,k \neq j=1}^{N} J_{-1jk} x_j x_k \right) \delta \left( \tilde{h}_{-10} - \sum_{k=1}^{N} J_{-10k} x_k \right). \quad (10)$$

We expect that $\tilde{h}_0$, $\tilde{h}_{-1}$, and $\tilde{h}_{-10}$ are the sums of independent random variables and hence we can assume that these three local fields obey a three-dimensional Gaussian distribution: $P_N(\tilde{h}_0, \tilde{h}_{-1}, \tilde{h}_{-10}) = \frac{1}{\sqrt{(2\pi)^N \det S}} \exp\{-\frac{1}{2} v^T S^{-1} v\}$, where $v^T = (\tilde{h}_0 - \langle \tilde{h}_0 \rangle_N, \tilde{h}_{-1} - \langle \tilde{h}_{-1} \rangle_N, \tilde{h}_{-10} - \langle \tilde{h}_{-10} \rangle_N)$ with $\langle \tilde{h}_0 \rangle_N$, $\langle \tilde{h}_{-1} \rangle_N$ and $\langle \tilde{h}_{-10} \rangle_N$ representing the $N$-body averages of the local fields and $S$ their covariance matrix

$$S = \begin{pmatrix} \sigma^2 & \sigma^2_{10} & \sigma^2_0 \left( -10 \right) \\ \sigma^2_{10} & \sigma^2 & \sigma^2 \left( -1 \right) \\ \sigma^2_0 \left( -10 \right) & \sigma^2 \left( -1 \right) & \hat{\sigma}^2 \end{pmatrix}. \quad (11)$$

Since the local fields are expected to be weakly correlated, one can expect that $\sigma^2_{10} = O(1/\sqrt{N})$, $\sigma^2_0 \left( -10 \right) = O(N^{-3/2})$ and $\sigma^2 \left( -1 \right) = O(N^{-3/2})$. Furthermore we find that $\hat{\sigma}^2 = O(1/N)$ by definition. Thus $S^{-1}$ reads

$$S^{-1} = \begin{pmatrix} \frac{1}{\sigma^2} & -\frac{\sigma^2_{10}}{\sigma^2} & -\frac{\sigma^2_0 \left( -10 \right)}{\sigma^2 \hat{\sigma}^2} \\ -\frac{\sigma^2_{10}}{\sigma^2} & \frac{1}{\sigma^2} & -\frac{\sigma^2 \left( -1 \right)}{\sigma^2 \hat{\sigma}^2} \\ -\frac{\sigma^2_0 \left( -10 \right)}{\sigma^2 \hat{\sigma}^2} & -\frac{\sigma^2 \left( -1 \right)}{\sigma^2 \hat{\sigma}^2} & \frac{1}{\hat{\sigma}^2} \end{pmatrix} + O \left( \frac{1}{N} \right). \quad (12)$$

32
Then the \((N+2)\)-body average of the state of the cavity neuron \(x_0\) is evaluated as

\[
\langle x_0 \rangle_{N+2} = C \int dx_0 dx_{-1} x_0 \left[ 1 + \left( \beta\langle \hat{h}_{-10} \rangle_N + \beta^2 \sigma_{-10}^2 \right) x_0 x_{-1} + O(1/N) \right] \\
\times e^{-\beta[\phi(x_0)+\phi(x_{-1})]} \exp \left\{ \beta\langle \hat{h}_0 \rangle_N x_0 + \beta\langle \hat{h}_{-1} \rangle_N x_{-1} + \frac{\beta^2 \sigma^2}{2} (x_0^2 + x_{-1}^2) \right\}
\]

up to \(O(1/\sqrt{N})\), since \(\langle \hat{h}_{-10} \rangle_N = O(1/\sqrt{N})\) by definition. The normalization constant \(C\) reads

\[
C^{-1} = \int dx_0 dx_{-1} \left[ 1 + \left( \beta\langle \hat{h}_{-10} \rangle_N + \beta^2 \sigma_{-10}^2 \right) x_0 x_{-1} + O(1/N) \right] \\
\times e^{-\beta[\phi(x_0)+\phi(x_{-1})]} \exp \left\{ \beta\langle \hat{h}_0 \rangle_N x_0 + \beta\langle \hat{h}_{-1} \rangle_N x_{-1} + \frac{\beta^2 \sigma^2}{2} (x_0^2 + x_{-1}^2) \right\}.
\]

Substituting eq.(14) into eq.(13) yields

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

where \(F\) is the transfer function which is defined as

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

and

\[
G(y) = \frac{\int dx x^2 \exp \left\{ -\beta\phi(x) + \beta y x + \frac{\beta^2 \sigma^2}{2} x^2 \right\}}{\int dx \exp \left\{ -\beta\phi(x) + \beta y x + \frac{\beta^2 \sigma^2}{2} x^2 \right\}}.
\]

Since \(\langle \hat{h}_{-10} \rangle_N = O(1/\sqrt{N})\) and \(\sigma_{-10}^2 = O(1/\sqrt{N})\), \(\langle x_0 \rangle_{N+2}\) is given as

\[
\langle x_0 \rangle_{N+2} = F \left( \langle \hat{h}_0 \rangle_N \right)
\]

in the limit of large \(N\).

Now our task is to express \(\langle \hat{h}_0 \rangle_N\) in terms of \(\langle \cdot \rangle_{N+2}\). To do this we calculate \(\langle \hat{h}_0 \rangle_{N+2}\) similarly to the above derivation of \(\langle x_0 \rangle_{N+2}\):

\[
\langle \hat{h}_0 \rangle_{N+2} = \hat{h}_0 + \beta \sigma^2 \langle x_0 \rangle_{N+2} + O \left( \frac{1}{\sqrt{N}} \right).
\]

Notice that the second term of the right hand side of eq.(19) is the Onsager reaction term. In our case we take two cavities, i.e., \(i = 0, -1\), however, the
Onsager reaction from site $i = -1$ to the site $i = 0$ is negligible in the limit of large $N$. Substituting eq.(19) into eq.(18) we find

$$\langle x_0 \rangle_{N+2} = F \left( \langle \hat{h}_0 \rangle_{N+2} - \beta \sigma^2 \langle x_0 \rangle_{N+2} \right).$$

(20)

Since the true local field $h_0$ is given by $h_0 = \hat{h}_0 + \hat{h}_{-10} x_{-1}$ and $\hat{h}_{-10} = O(1/\sqrt{N})$, we can replace $\langle \hat{h}_0 \rangle_{N+2}$ in eq.(20) with the true local field $\langle h_0 \rangle_{N+2}$ to have

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

(21)

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 states of neurons $\langle x_j x_k \rangle_{N+2}$. In order to find the expression of the local field $\langle h_i \rangle_{N+2}$ in terms of the thermal averages of states of neurons, it will be necessary to employ 2-cavity method as seen below:

$$\langle x_0 x_{-1} \rangle_{N+2} = F \left( \langle \hat{h}_0 \rangle_{N} \right) F \left( \langle \hat{h}_{-1} \rangle_{N} \right) + \left( \beta \langle \hat{h}_{-10} \rangle_{N} + \beta^2 \sigma_{-10}^2 \right) F \left( \langle \hat{h}_0 \rangle_{N} \right) F \left( \langle \hat{h}_{-1} \rangle_{N} \right) + O(1/N).$$

(22)

Comparing this relation with eq.(15) and using $\langle \hat{h}_{-10} \rangle_{N+2} = \langle \hat{h}_{-10} \rangle_{N} + O(1/N)$ and $\langle x_0^2 \rangle_{N+2} = G(\langle \hat{h}_0 \rangle_{N}) + O(1/\sqrt{N})$, one obtains

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

(23)

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

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

(24)

where $U$ is the average of susceptibilities $U = \frac{\beta}{N+2} \sum_j \langle x_j^2 \rangle_{N+2} - \langle x_j \rangle_{N+2}^2$.

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 site index, on the other hand $\sigma_{-10}^2$ does not affect the local field. Eq.(24) together with eq.(21) yields the pre-TAP equations [10] for $p = 3$ Hopfield model:

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

(25)

with $\Gamma_{\text{TAP}} = \beta \sigma^2 - \frac{\alpha U^2}{2 \beta}$, 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.(11) has still to be determined to obtain the TAP equation.
Eq. (25) defines a deterministic analog network of associative memory with the transfer function $F$ corresponding to the original stochastic analog neural network (9). To obtain $\sigma^2$ together with the order parameter equations we then apply the SCSNA to eq. (25). Thus we can apply the SCSNA to the original stochastic analog network as seen in the previous section.

Substituting the expression of the local field (2) into the pre-TAP equation (25) and comparing with eq. (18), it follows that $\Gamma_{SCSNA} = \Gamma_{TAP}$ [10], since $\langle \tilde{h}_i \rangle N - \frac{1}{2} \xi^1 (m^1)^2$ obeys a Gaussian random variable. Thus we find

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

where $\Gamma_{TAP}$ is given by eq. (6) and determines the shape of the transfer function $F$. Then we finally obtain the TAP equation as

$$\langle x_i \rangle_N = F \left( \sum_{j<k(\neq i)} J_{ijk} \langle x_j \rangle_N \langle x_k \rangle_N - \Gamma_{SCSNA} \langle x_i \rangle_N \right) \quad (27)$$

with $\Gamma_{SCSNA}$ and the form of the transfer function $F$ self-consistently determined by the above Eqs. (5), (6), (7), (8) and (26).

3 Pattern Adding Approach to Onsager Reaction Term

In the previous section, we have evaluated the coefficient of the Onsager reaction term characteristic to the TAP equation using the SCSNA. As we have seen in the previous section, the Onsager reaction term consists of the variance of the local field $\sigma^2$ and the cross-correlation between states of neurons. In our case, we have directly evaluated the cross-correlation using the 2-cavity method. Furthermore, the variance of the local field has been obtained by the SCSNA with the relation eq. (26). In this section, we will perform the alternative method to have the variance of the local field $\sigma^2$.

In the previous section, we found the relation $\Gamma_{SCSNA} + \frac{\alpha U^2}{2 \beta} - \beta \sigma^2 = 0$ using the SCSNA. Since $\Gamma_{SCSNA} = \alpha U q$ and $U = (\hat{q} - q)/\beta$ where $\hat{q} = \frac{1}{N} \sum_i \langle x_i \rangle^2$ and the Edward-Anderson order parameter $q = \frac{1}{N} \sum_i \langle x_i \rangle^2$, the variance of the local field $\sigma^2$ can be expressed as $\beta \sigma^2 = \frac{\alpha}{\beta} (\hat{q}^2 - q^2)$ by using the relation $\frac{U}{\beta} = \frac{1}{N} \sum_i (\langle x_i^2 \rangle - \langle x_i \rangle^2)$. In this section we will show that the variance of the local field can also be obtained by adding a new random memory pattern to the network [2, 10].

The variance of the local field is given as $\sigma^2 = \langle (h_0 - \langle h_0 \rangle_N)^2 \rangle_N$ by definition. Substituting the Hebb learning rule into this definition of the variance of the local field, we have

$$\sigma^2 = \frac{1}{4} \sum_{\mu} \left( \langle m^\mu \rangle^4 \right)_N - \left( \langle m^\mu \rangle^2 \right)_N^2$$

35
\[
\begin{align*}
&+ \frac{1}{4} \sum_{\mu \neq \nu} \xi_0^\mu \xi_0^\nu \left( \left( \langle m^\mu \rangle^2 \langle m^\nu \rangle^2 \right)_N - \left( \langle m^\mu \rangle \right)_N \left( \langle m^\nu \rangle \right)_N \right) \\
&- \alpha \frac{1}{4N^2} \sum_{i,j} \left( \langle x_i^2 x_j^2 \rangle_N - \langle x_i^2 \rangle_N \langle x_j^2 \rangle_N \right) \\
&- \frac{1}{4N^2} \sum_{\mu \neq \nu} \sum_{i,j} \xi_0^\mu \xi_0^\nu \left( \langle x_i^2 x_j^2 \rangle_N - \langle x_i^2 \rangle_N \langle x_j^2 \rangle_N \right).
\end{align*}
\]

Since one can expect that the neurons in the system under consideration are weakly coupled, it is expected that \( \langle x_i^2 x_j^2 \rangle_N - \langle x_i^2 \rangle_N \langle x_j^2 \rangle_N = O(1/\sqrt{N}) \) for \( i \neq j \). Thus the orders of the third and fourth terms of eq.(28) are \( O(1/\sqrt{N}) \). Furthermore the second term of eq.(28) is also negligible since the overlap \( m^\mu \) does not include the pattern \( \xi_0^\mu \) at the cavity site. Therefore the variance of the local field \( \sigma^2 \) can be expressed as \( \sigma^2 = \frac{1}{4} \sum_\mu \left( \langle (m^\mu)^4 \rangle_N - \langle (m^\mu)^2 \rangle_N^2 \right) + O(1/\sqrt{N}) \). To obtain \( \langle (m^\mu)^4 \rangle_N \) and \( \langle (m^\mu)^2 \rangle_N \), we increase the number of embedded patterns by one, i.e., \( \tilde{p} \rightarrow \tilde{p} + 1 \).

The Hamiltonian of \( \tilde{p} \)-pattern system is given as \( H_\tilde{p} = \sum_{i=1}^N \phi(x_i) - \sum_{i<j<k} J_{ijk}^\tilde{p} x_i x_j x_k \), where \( J_{ijk}^\tilde{p} \) is the synaptic coupling weight tensor of \( \tilde{p} \)-pattern system. Using the Hebb learning rule, the coupling tensor of \( (\tilde{p} + 1) \)-pattern system yields \( J_{ijk}^{\tilde{p}+1} = J_{ijk}^\tilde{p} + \frac{1}{N^2} \sum_{i<j<k} (\xi_0^0)^3 \sum_{\delta} \), where \( 0 \)th pattern is added. Then the Hamiltonian of the \( (\tilde{p} + 1) \)-pattern system can be expressed in terms of the overlap of the new added pattern \( m^0 \) as

\[
H_{\tilde{p}+1} = H_\tilde{p} - \frac{N}{6} \langle m^0 \rangle^3 + \frac{1}{2} \left( \frac{1}{N} \sum_{i=1}^N x_i^2 \right) m^0 - \frac{1}{3N^2} \sum_{i=1}^N \xi_0^0 x_i^3.
\]

The last term is negligible since this term is \( O(N^{-3/2}) \). Since \( m^0 \) is \( O(1/\sqrt{N}) \), one introduces a new variable \( \tilde{m}_0 = \sqrt{N} m^0 \) to extract the effect of the \( 0 \)th pattern on the system. Using this new variable the Hamiltonian of the \( (\tilde{p} + 1) \)-pattern system (29) is rewritten as \( H_{\tilde{p}+1} = H_\tilde{p} - \frac{\tilde{m}_0^3}{6\sqrt{N}} + \frac{\tilde{m}_0}{2\sqrt{N}} + O(N^{-3/2}) \), where \( \tilde{q} = \frac{1}{N} \sum_{i=1}^N x_i^2 \). Then the probability density of \( \tilde{m}_0 \) is given as

\[
P_{\tilde{p}+1} (\tilde{m}_0) = \frac{Z_{\tilde{p}+1}}{Z_{\tilde{p}}} \exp \left\{ \frac{\beta \tilde{m}_0 (\tilde{m}_0^2 - 3\tilde{q})}{6\sqrt{N}} \right\} \times \frac{1}{Z_{\tilde{p}}} \int \prod_{i=1}^N dx_i \delta \left( \tilde{m}_0 - \sum_{i=1}^N \frac{\xi_0^0 x_i}{\sqrt{N}} \right) e^{-\beta H_\tilde{p}},
\]

where \( Z_{\tilde{p}+1} \) and \( Z_{\tilde{p}} \) are the partition functions of the \( \tilde{p} \)- and \( (\tilde{p} + 1) \)-pattern system respectively. Here we assume that the probability distribution function of \( \tilde{m}_0 \) in the \( \tilde{p} \)-pattern system \( P_{\tilde{p}} (\tilde{m}_0) = \frac{1}{Z_{\tilde{p}}} \int \prod_{i=1}^N dx_i \delta (\tilde{m}_0 - \frac{1}{\sqrt{N}} \sum_{i=1}^N \xi_0^0 x_i) e^{-\beta H_\tilde{p}} \) obeys a Gaussian distribution. Then the probability density of \( \tilde{m}_0 \) in the \( (\tilde{p} + 1)\)-
pattern system reads

\[ P_{\tilde{p}+1}(\tilde{m}_0) = \frac{Z_{\tilde{p}}}{\sqrt{2\pi U Z_{\tilde{p}+1}}} \exp \left( \frac{\beta \tilde{m}_0^3 - 3\beta \tilde{q} \tilde{m}_0}{6\sqrt{N}} - \frac{(\tilde{m}_0 - \langle \tilde{m}_0 \rangle_{\tilde{p}})^2}{2U} \right), \]  

(31)

where \( \tilde{U} \) is the variance of the Gaussian distribution and \( \langle \tilde{m}_0 \rangle_{\tilde{p}} \) is the average of \( \tilde{m}_0 \) in the \( \tilde{p} \)-pattern system. The variance \( \tilde{U} \) is evaluated as

\[ \tilde{U} = \langle (\delta \tilde{m}_0)^2 \rangle_{\tilde{p}} = \frac{1}{N} \sum_{i=1}^{N} \langle (\delta x_i)^2 \rangle_{\tilde{p}} + O \left( \frac{1}{\sqrt{N}} \right). \]  

(32)

Since \( \langle (\delta x_i)^2 \rangle_{\tilde{p}} \) is the susceptibility of the state of neuron \( i \), eq.(32) can be expressed as \( \tilde{U} = \frac{U}{4} + O(1/\sqrt{N}) \), where \( U \) is defined in the previous section. Thus the probability density of \( \tilde{m}_0 \) in the \((\tilde{p}+1)\)-pattern system (30) is expressed as

\[ P_{\tilde{p}+1}(\tilde{m}_0) = \frac{\tilde{C}}{\sqrt{2\pi U/\beta}} \exp \left\{ \frac{\beta \tilde{m}_0^3}{6\sqrt{N}} - \frac{(\tilde{m}_0 - \langle \tilde{m}_0 \rangle_{\tilde{p}})^2}{2U/\beta} \right\} \]  

(33)

up to \( O(1/\sqrt{N}) \), where \( \tilde{C} \) is a normalization constant. Integrating the probability density of \( \tilde{m}_0 \) in the \((\tilde{p}+1)\)-pattern system (33), one obtains the perturbation expression of the normalization constant \( \tilde{C} \) as \( C^{-1} = 1 + \frac{\tilde{q}}{6\sqrt{N}} \frac{\beta}{\beta + \tilde{q}} \langle \tilde{m}_0 \rangle_{\tilde{p}} + \langle \tilde{m}_0 \rangle_{\tilde{p}}^3 - \frac{\tilde{q}}{2\sqrt{N}} \langle \tilde{m}_0 \rangle_{\tilde{p}} + O(1/N) \). Then we have

\[ \langle \tilde{m}_0^2 \rangle_{\tilde{p}+1} = \frac{U}{\beta} + \langle \tilde{m}_0 \rangle_{\tilde{p}}^2 + \frac{1}{\sqrt{N}} \left\{ \frac{2U^2}{\beta} \langle \tilde{m}_0 \rangle_{\tilde{p}} + U \langle \tilde{m}_0 \rangle_{\tilde{p}}^3 - \tilde{q} U \langle \tilde{m}_0 \rangle_{\tilde{p}} \right\} \]  

(34)

and

\[ \langle \tilde{m}_0^4 \rangle_{\tilde{p}+1} = \frac{3U^2}{\beta^2} + \frac{6U}{\beta} \langle \tilde{m}_0 \rangle_{\tilde{p}}^2 + \langle \tilde{m}_0 \rangle_{\tilde{p}}^4 \]  

\[ + \frac{1}{\sqrt{N}} \left\{ \frac{16U^3}{\beta^3} \langle \tilde{m}_0 \rangle_{\tilde{p}} + \frac{14U^2}{\beta} \langle \tilde{m}_0 \rangle_{\tilde{p}}^3 + 2U \langle \tilde{m}_0 \rangle_{\tilde{p}}^5 \right\} \]  

\[ - \frac{\tilde{q}}{\sqrt{N}} \left\{ \frac{6U^2}{\beta} \langle \tilde{m}_0 \rangle_{\tilde{p}} + 2U \langle \tilde{m}_0 \rangle_{\tilde{p}}^3 \right\} \]  

(35)

up to \( O(1/\sqrt{N}) \). Using these equations (34) and (35), we obtain the variance of the local field \( \sigma^2 \) as

\[ \sigma^2 = \frac{\alpha U^2}{2\beta^2} + \frac{U}{\beta N^2} \sum_{\mu} \langle \tilde{m}_\mu \rangle_{\tilde{p}}^2 + O \left( 1/\sqrt{N} \right). \]  

(36)

Noticing that \( \sum_{\mu=1}^{N^2} \langle \tilde{m}_\mu \rangle_{\tilde{p}}^2 = \alpha N^2 q + O(N) \), eq.(36) can be rewritten as \( \sigma^2 = \frac{\alpha U^2}{2\beta^2} + \frac{\alpha U q}{\beta} + O(1/\sqrt{N}) \). Since the relation \( U/\beta = \tilde{q} - q \) holds by the definition
of $U$, we finally obtain $\sigma^2 = \frac{q}{2} \left( \bar{q}^2 - q^2 \right)$. This result is consistent with the calculation of the SCSNA. Since the variance of the local field $\sigma^2$ can be obtained directly using the pattern adding method above mentioned, the cavity method where only one cavity is taken except the 2-cavity method we performed in the previous section can reproduce the TAP equation resorting the SCSNA with the pattern addition.

4 Summary and Conclusion

We have taken advantage of the close relationship between the TAP equation and the SCSNA for analog networks with 3-body interaction based on the Hebb learning rule to compute the Onsager reaction term appearing in the TAP equation together with the set of order parameter equations. Characteristic to the TAP equation for the case of such interaction is that the Onsager reaction term consists of the two terms: 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 recipe of 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. Furthermore, 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 deterministic analog networks having 3-body interaction and transfer function $F$ of arbitrary shape irrespective of whether it is monotonic or not [5, 16–19]. The order parameter equations obtained in the framework of the SCSNA reproduce those given by the replica method.

In section 3, we have seen that adding a random memory pattern to the network has recovered the result of the SCSNA. More precisely, using this conventional recipe we have reproduced the variance of the local field which determines the form of the transfer function. Characteristic to the computation of the coefficient of the Onsager reaction term for the stochastic analog network with many-body interaction is the need for obtaining both the variance of the local field and the cross-correlation between the states of neurons, although the coefficient of the Onsager reaction term for the Hopfield model with 2-body interaction consists only of the variance of the local field. In section 3, we have seen that the variance of the local field could be directly obtained by the pattern addition. However, it is incomplete for obtaining the TAP equation of such network to apply the conventional cavity method where only one cavity is taken with the pattern addition. On the other hand, the coefficient of the Onsager reaction term itself has been directly obtained in the framework of the SCSNA as we have seen in section 2. Since the cross-correlation has been computed using the 2-cavity method as we have seen in section 2, the SCSNA which gives the coefficient of the Onsager reaction term directly connects the concepts of the
cavity and pattern adding method. 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.

References