A continuous-time model of autoassociative neural memories utilizing the noise-subspace dynamics

(Received ..... ; Accepted in final form .....)

Abstract. This paper presents a continuous-time model of Autoassociative Neural Memories (ANMs) which correspond to a modified version of pseudoinverse-type ANMs. This ANM model is derived from minimizing the energy function for a modular neural network. Through the eigendecomposition of the connection matrix, we show that the dynamical properties of the ANM are qualitatively different in the two state subspaces: a pattern-subspace and a noise-subspace. The proposed ANM has a distinctive feature in the noise-subspace dynamics. The size of basins of attraction can be varied by controlling the contribution of the noise-subspace dynamics to the whole network. The first simulation confirms this attractive feature. In the second simulation, we investigate the performance robustness of the ANM for several kinds of correlated pattern sets. These simulation results confirm the usefulness of the proposed ANM.

Key words: associative memory, basins of attraction, continuous-time dynamics, correlated patterns, Hopfield network, modular neural network

Abbreviations: ANM – autoassociative neural memory; CCHN – Cross-Coupled Hopfield Nets; HK – Ho-Kashyap; PANM – pseudoinverse-type ANM; DB-PANM – diagonally biased pseudoinverse-type ANM

1. Introduction

The model of association and memory is one of the most fundamental and important research subjects in the field of artificial neural networks. Actually, a large number of successful studies on associative neural memories have been done so far. The models can be classified into various groups such as autoassociative or heteroassociative recollection, single-pass or iterative association (the latter is subdivided into associations with discrete-time or continuous-time dynamics), discrete-valued (binary or bipolar) or continuous-valued outputs, and so forth.

The performance of associative neural memories is largely related to the way of synthesizing their connection matrices. We can roughly classify into three types of connection matrices: correlation-type [1, 2], pseudoinverse-type [2, 3] and Ho-Kashyap(HK)-type [4, 5]. The correlation-type associative neural memories are the most popular approach in which connection matrices are given by the outer-product of memory patterns. These are very simple to implement, but the poor memory capacity and the intolerance of noise have been pointed out. The pseudoinverse-type associative neural memories, whose connection matrices are synthesized by using the pseudoinverse of the pattern ma-

trices, outperform the correlation-type ones in the memory capacity. They are useful especially for the memorization of correlated patterns, but the performance in noise tolerance is not very good [6]. It is well known that their connection matrices correspond to a minimum mean squared error solution that is obtained from minimizing the total errors between network outputs and memory patterns to be retrieved. The HK-type associative neural memories are based on the algorithm proposed by Ho and Kashyap [7] which provides the solution for a system of linear inequalities. It has been reported that they can also realize the high association performance [4, 5].

This paper presents a new connection synthesis which consequently corresponds to a modified version of pseudoinverse-type Autoassociative Neural Memories (ANMs) with continuous-valued continuous-time dynamics. This method is derived from minimizing the two types of energy functions which have been originally defined for a modular neural network called Cross-Coupled Hopfield Net [8]. As will be stated later, the performance improvement in this method is originated in the operation that bias values are added to the diagonal elements of the conventional connection matrix (we call such a connection synthesis "biased method").

In Section 2, the derivation of the ANM model is described, then it is compared with the Kanter's model [6] and the Gorodnichy's model [9] in which similar biased methods are applied to the conventional pseudoinverse-type ANM. Furthermore, we discuss the dynamical properties of the proposed ANM through the eigendecomposition of the connection matrix. This consideration clarifies the distinction in the dynamical characteristics between the conventional pseudoinverse-type ANM and the proposed ANM. We also show that this distinctive feature in the proposed ANM realizes the variable basins of attraction and the high association performance for correlated pattern sets. In Section 3, we confirm these properties through some computer simulations. Section 4 presents a conclusion.

2. A Continuous-Time Model of Autoassociative Neural Memories

2.1. The Derivation

The ANM model we will propose here is obtained from minimizing the energy function which has been defined for a modular neural network called Cross-Coupled Hopfield Nets (CCHN). The CCHN is composed of plural Hopfield networks (HNs) which are mutually connected via

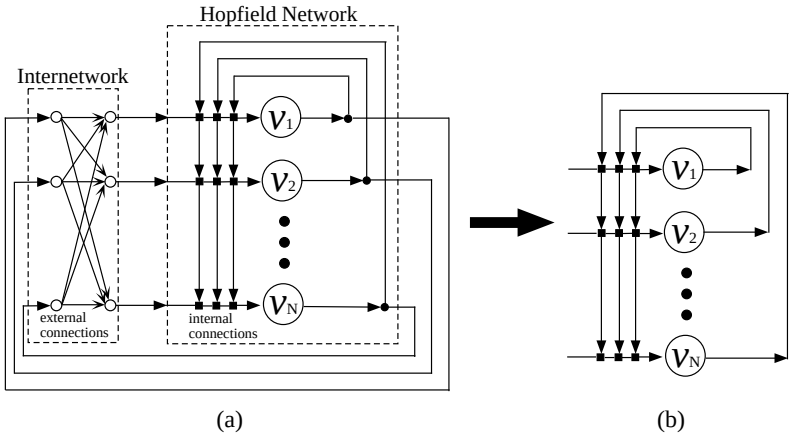

Figure 1. (a) A schematic diagram of the connections in a single-module Cross-Coupled Hopfield Nets (CCHN) with two-layer internetwork. (b) This single-module CCHN is degenerated into a simple Hopfield network as a result of the model-derivation. The white circles mean network units and the black diamonds symbolize connections of the Hopfield network.

feedforward networks called internetworks [8, 10, 11]. The ANM model is derived from the simplest version of the CCHN: a single-module CCHN with two-layer internetwork (see Figure 1(a)). To be more precise, this CCHN consists of a single HN which has the external connections realized by a linear-mapping internetwork as well as its own internal connections. As shown in Figure 1, these two types of connections can be degenerated into a single connection matrix of a Hopfield network through the following model-derivation.

Suppose that the HN in a single-module CCHN has N units and its state vector is denoted by $\boldsymbol{v} = (v_1, \cdots, v_N)'$, where $'$ means the transposition of vectors and matrices. Then, we can define the following energy function to be minimized:

$$E = -\frac{\alpha}{2}\boldsymbol{v}'\boldsymbol{T}\boldsymbol{v} + \frac{\beta}{2}\mathrm{tr}(\boldsymbol{W}\boldsymbol{W}') + \frac{1}{2}(\boldsymbol{v} - \boldsymbol{W}\boldsymbol{v})'(\boldsymbol{v} - \boldsymbol{W}\boldsymbol{v}), \qquad (1)$$

where $\mathrm{tr}(\cdot)$ means the trace of matrices. $\boldsymbol{T} \in \boldsymbol{R}^{N\times N}$ and $\boldsymbol{W} \in \boldsymbol{R}^{N\times N}$ are the connection matrices of the HN and the internetwork, respectively. α and β are positive constants which determine the relative importance of the three terms on the right-hand side of Equation (1). Note that the internetwork output is represented by $\boldsymbol{W}\boldsymbol{v}$ in the third term. The first term in Equation (1) is equivalent to the energy function for the HN, while the third term is the energy function evaluating

the errors between the HN outputs and the internetwork outputs[1]. The second term is an energy function that prevents the internetwork connections $\boldsymbol{W}$ from going to infinity.

To be implemented as ANMs, we have to determine the network dynamics such that the value of the above energy function always decreases and its minimum points correspond to M memory patterns $\boldsymbol{\xi}^{\mu}(\mu = 1, \cdots, M)$. First, we calculate the time derivative of E. Suppose that $\boldsymbol{T}$ is a symmetrical constant matrix, then the following derivative is obtained:

$$\begin{aligned}\frac{dE}{dt} &= -\alpha \frac{d\boldsymbol{v}'}{dt}\boldsymbol{T}\boldsymbol{v} + \frac{d\boldsymbol{v}'}{dt}(\boldsymbol{I}-\boldsymbol{W})'(\boldsymbol{I}-\boldsymbol{W})\boldsymbol{v} \\ &\quad +\beta \cdot \mathrm{tr}(\boldsymbol{W}\frac{d\boldsymbol{W}'}{dt}) - \boldsymbol{v}'\frac{d\boldsymbol{W}'}{dt}(\boldsymbol{I}-\boldsymbol{W})\boldsymbol{v}, \end{aligned} \tag{2}$$

where $\boldsymbol{I}$ is a $N \times N$ identity matrix. To reduce the energy, the above dE/dt should be less than or equal to zero. Thus, we can determine the following network dynamics (see Appendix A):

$$c\frac{d\boldsymbol{u}}{dt} = \{\alpha\boldsymbol{T} - (\boldsymbol{I}-\boldsymbol{W})'(\boldsymbol{I}-\boldsymbol{W})\}\boldsymbol{v} \tag{3}$$

$$\eta\frac{d\boldsymbol{W}}{dt} = -\boldsymbol{W} + \frac{1}{\beta}(\boldsymbol{I}-\boldsymbol{W})\boldsymbol{v}\boldsymbol{v}', \tag{4}$$

where

$$\boldsymbol{v} = \tanh(k\boldsymbol{u}). \tag{5}$$

Here, $\boldsymbol{u}$ is the internal potential vector of a unit which is transferred to the network output $\boldsymbol{v}$ by a nonlinear function in Equation (5). c and η are positive time constants, k is a positive steepness parameter. Equation (3) corresponds to an update rule of network states (i.e. the activity dynamics), and Equation (4) gives a learning rule of internetwork weights (i.e. the connection dynamics).

Although the above differential equations should be solved simultaneously in the strict sense of word, we solve them independently in order to make it easy to ensure the state convergence to memory patterns. This is realized under the following assumptions: (1) $\boldsymbol{W}$ is modified very slowly as compared with the variation of unit potential $\boldsymbol{u}$ (i.e. $c \ll \eta$ holds for time constants), (2) memory patterns $\boldsymbol{\xi}^{\mu}(\mu = 1, \cdots, M)$ are repeatedly presented to a network, and the variation of $\boldsymbol{W}$ at one

[1] In the context of modular neural networks, these terms are refereed as the energy functions for modules and their interactions. In this paper, we don't care about why such functions should be defined (see [11] for details).

time is small enough. Under these assumptions, Equation (4) is approximately rewritten as follows:

$$\eta \frac{d\boldsymbol{W}}{dt} = -\boldsymbol{W} + \frac{1}{\beta}(\boldsymbol{I} - \boldsymbol{W})\boldsymbol{X}\boldsymbol{X}', \tag{6}$$

where

$$\boldsymbol{X} = (\boldsymbol{\xi}^1 \ \boldsymbol{\xi}^2 \ \cdots \ \boldsymbol{\xi}^M). \tag{7}$$

It is well known that if β is small enough, the solution of $\boldsymbol{W}$ is approximately given by $\boldsymbol{X}\boldsymbol{X}^+$ in which $\boldsymbol{X}^+$ is the Moore-Penrose pseudoinverse of $\boldsymbol{X}$, i.e. $\boldsymbol{X}^+ = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$.

Since the first energy term in Equation (1) should be minimized when $\boldsymbol{v} = \boldsymbol{\xi}^\mu$, we can adopt any $\boldsymbol{T}$ as long as $\boldsymbol{T}\boldsymbol{\xi}^\mu = \boldsymbol{\xi}^\mu$ holds for any μ. Let us also adopt $\boldsymbol{X}\boldsymbol{X}^+$ for $\boldsymbol{T}$. Then, Equation (3) is rewritten as follows:

$$c\frac{d\boldsymbol{u}}{dt} = \{\alpha \boldsymbol{X}\boldsymbol{X}^+ - (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^+)'(\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^+)\}\boldsymbol{v}. \tag{8}$$

Noting that $\boldsymbol{X}\boldsymbol{X}^+$ is symmetric, the following approximation for the activation dynamics is obtained:

$$c\frac{d\boldsymbol{u}}{dt} = \{(1+\alpha)\boldsymbol{X}\boldsymbol{X}^+ - \boldsymbol{I}\}\boldsymbol{v} \quad (c > 0). \tag{9}$$

This differential equation corresponds to the ANM we propose here. Considering that the dynamics of the conventional pseudoinverse-type ANM is presented by Equation (10), the primary difference between the proposed ANM and the conventional one lies in the diagonal elements of their connection matrices.

$$c'\frac{d\boldsymbol{u}}{dt} = \boldsymbol{X}\boldsymbol{X}^+\boldsymbol{v} \quad (c' > 0) \tag{10}$$

For the notational convenience, the proposed model is denoted by DB-PANM (Diagonally Biased Pseudoinverse-type ANM) in the sense that bias values are added to the diagonal elements of connection matrices (we call such a connection synthesis "biased method").

Some other biased methods have been proposed by Kanter et al. [6], Gorodnichy et al. [9] and Ikeda [12]. In the Gorodnichy's model, the diagonal elements W_{ii} are given by the following equation:

$$W_{ii} = D \cdot W_{ii}^{org}, \tag{11}$$

where W_{ii}^{org} is the ith diagonal element of the original connection matrix and D $(0 < D < 1)$ is a positive constant called "desaturating coefficient". Since the connection matrix of the Kanter's model has zero

diagonal elements, this model can be considered as a special case of the Gorodnichy's model. The DB-PANM is equivalent to the Gorodnichy's model only when the diagonal elements are all the same. It is known that this situation is realized when the number of memory patterns is large enough [13]. Therefore, one can say that the differences between the Gorodnichy's model and the DB-PANM are distinctive when the number of memory patterns are not too large (another difference is that the Gorodnichy's model is proposed as a bipolar-valued discrete-time model).

The connection matrix of the Ikeda's model is very similar to that of the DB-PANM; however, the Ikeda's model was proposed as an associative memory model with bipolar-valued discrete-time dynamics [12]. The Ikeda's dynamics was presented as equivalent dynamics of the conventional pseudoinverse-type ANM with hysteresis neurons, therefore the motivation of the two models is quite different. We should notice that there is one big difference in the stability between the two models. In the DB-PANM, the asymptotic stability is always ensured for any pattern sets (note that the dynamics is derived such that the energy function value monotonically reduces), while the Ikeda's model can converge to oscillatory states as well as fixed points.

2.2. The Dynamical Nature

We shall discuss the dynamical nature of the DB-PANM through the eigendecomposition of the connection matrix. Using the eigenvectors of $\boldsymbol{X}\boldsymbol{X}'$, Equation (9) can be further rewritten as follows:

$$c\frac{d\boldsymbol{u}}{dt} = \boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{v}, \tag{12}$$

where

$$\boldsymbol{A} = diag\,(\underbrace{\alpha\ \cdots\ \alpha}_{r}\ \underbrace{-1\ \cdots\ -1}_{N-r})' \in \mathcal{R}^{N\times N} \quad (\alpha > 0). \tag{13}$$

r is the rank of $\boldsymbol{X}\boldsymbol{X}'$. The column vectors of $\boldsymbol{Q}$ correspond to the eigenvectors of $\boldsymbol{X}\boldsymbol{X}'$ which are arranged by magnitude of the corresponding eigenvalues (i.e. the first column vector has the largest eigenvalue). It is well known that the pattern vectors span the r-dimensional subspace whose orthogonal bases are the first r column vectors of $\boldsymbol{Q}$. This subspace is called "pattern-subspace", while its orthogonal complementary subspace is called "noise-subspace". Since the connection matrices of the conventional pseudoinverse-type ANM and the DB-PANM are diagonalized by the same eigenvectors, the dynamical properties of these

ANMs can be discussed in the same eigenspaces (i.e. pattern-subspace and noise-subspace).

The change of network states is determined by two types of projection: the linear projection of a network input via feedback connections (see the right-hand side of Equation (12)) and the nonlinear projection by a transfer function (see Equation (5)). As well known in ANMs, the nonlinear projection is important for the exact recall of binary (or bipolar) patterns. This projection, however, only gives the quantization of vector elements by which the internal potential vector $\boldsymbol{u}$ is mapped to the nearest hypercube corner of the state space. Therefore, it seems reasonable to say that the state change in ANMs is primarily determined by the linear projection [14].

As seen in Equations (12) and (13), the linear projection in the DB-PANM is featured by two different eigenvalues: α for the pattern-subspace and -1 for the noise-subspace. This fact suggests that the state components in the pattern-subspace are monotonously augmented by the linear projection; on the contrary, those in the noise-subspace are monotonously reduced. Since the noise-subspace components are unnecessary for correct recollections, one can say that the noise-subspace dynamics in the DB-PANM realizes the noise reduction for corrupted inputs. On the other hand, it is well known that the eigenvalues for the conventional pseudoinverse-type ANM are $+1$ for the pattern-subspace and 0 for the noise-subspace. In this case, the noise-subspace components are not varied by the linear projection. It follows from the above considerations that the essential difference between the dynamical natures of the conventional pseudoinverse-type ANM and the DB-PANM lies in their noise-subspace dynamics.

Another feature of the DB-PANM is that the property of the linear projection can be changed by α values. If $\alpha < 1$, the noise-subspace dynamics dominates the pattern-subspace dynamics; that is, the state change is strongly influenced by the linear projection in the noise-subspace. If $\alpha > 1$, the reverse is true. This feature may give us various associations in the DB-PANM.

3. Simulations

3.1. Preliminary

Pseudoinverse-type ANMs have been often applied to the memorization of correlated patterns. Here, we suppose that memory patterns ξ^μ ($\mu = 1, \cdots, M$) are represented by N-dimensional bipolar vectors (i.e. $\xi^\mu \in$

$\{-1, 1\}^N$). Note that N represents the number of network units as well as the dimensions of memory pattern vectors.

In real world problems, memory patterns mutually correlate with others in various ways; some patterns are strongly correlated, but others are not so much. Therefore, when we investigate the association performance of a certain ANM, it is important to evaluate the performance for various kinds of correlated patterns as much as possible. Needless to say, the larger number of pattern sets are applied to the performance evaluation of an ANM, the higher reliability for the evaluation is ensured. To make the performance evaluation efficiently, we have proposed a randomly-generated cluster pattern set [15]. In this pattern set, we suppose that memory patterns are divided into some clusters and the patterns in the same cluster are mutually correlated at a given level. Such pattern sets are defined by the following parameters: the number of clusters C, the number of patterns L in a cluster, the average correlation $\bar{\sigma}$ between patterns in the same cluster, and the average correlation $\bar{\kappa}$ between cluster centroids.

In the simulations, we shall evaluate the association performance of DB-PANM for two types correlated pattern sets: one is the cluster pattern set described the above, the other is the alphabet pattern set. These pattern sets are generated in the following way:

Type-I (cluster pattern set) We use the cluster pattern sets whose parameters L and $\bar{\kappa}$ are fixed at the constant values: $L = 5$ and $\bar{\kappa} \simeq 0$. On the other hand, the parameters C and $\bar{\sigma}$ are varied in order to generate variations of the cluster pattern sets. Note that since L is fixed, the total number of memory patterns M is varied by C. Each memory pattern is represented by a 100 dimensional bipolar-valued vector (i.e. N=100). In order to avoid biased evaluations, 5 different memory pattern sets are generated and evaluated for each experiment.

Type-II (alphabet pattern set) We use a set of 26 alphabet patterns ("a" to "z") in which every pattern is represented by a $64(= 8 \times 8)$ dimensional bipolar-valued vector. The mean correlation among these patterns is 0.377, and the maximum and minimum correlations are 0.906 and 0.0, respectively.

After the connection matrix of a network is synthesized, we estimate the association performance through many trials of recollection. In each trial, a test pattern $\boldsymbol{v}_k$ is set to the network as its initial state, then the recall gets started. If the stable state is identical with the pattern vector $\boldsymbol{\xi}_r$ to be recalled, we say the trial succeeds. The state changes of the network are calculated based on the continuous-time continuous-valued

dynamics in Equation (9). Note that the effectiveness of the noise-subspace dynamics mentioned in the previous section is valid when the linear projection via feedback connections gives the time-differentiation (or time-difference) of unit states, i.e. the network dynamics should be represented as differential (or difference) equations.

The difficulty of trials is evaluated based on the following initial direction cosine d:

$$d = \frac{1}{N} \boldsymbol{v}_p' \boldsymbol{\xi}_r. \tag{14}$$

The small d means that an initial state of a network is far from the pattern to be retrieved, hence the trial is difficult to succeed. In this context, as a measure of the association performance, we adopt the critical direction cosine d_c which is defined by the minimum value of d when trials succeed over 98% probability. Needless to say, the smaller d_c means the greater basins of attraction on average. To estimate the d_c, we test the trials with a fixed set of 750 randomly generated test patterns for every memory pattern.

3.2. The performance for different α

In the first part of the simulations, we shall study the influence of α to the performance of the DB-PANM.

Three kinds of Type-I pattern sets are used which are different in the number of clusters: $C = 6, 10, 14$. The other parameters are as follows: $L = 5$, $\bar{\kappa} \approx 0$, and $\bar{\sigma} = 0.4$. Figure 2 shows the simulation results for different α. One can say that critical direction cosine d_c tends to be larger (i.e. the basins of attraction become smaller) when α becomes larger. Figure 2 also suggests that the optimal value of α is about 0.125 in all cases. Since the place where $\alpha = 1$ is the critical point that classifies the dynamical nature of the DB-PANM into two different types, the results mean that the predominance of the noise-subspace dynamics over the pattern-subspace dynamics contributes to the extension of basins of attraction. The results also suggest that it is possible to control the size of attractive basins to some extent by changing the α values.

3.3. The performance for different pattern sets

Next, we evaluate the association performance of the DB-PANM for several kinds of Type-I pattern sets which are different in the number of clusters C and the average pattern correlation $\bar{\sigma}$. We also evaluate the association performance for Type-II pattern sets: 26 alphabet patterns. The purpose of these experiments is to examine the robustness of the association performance for different pattern sets.

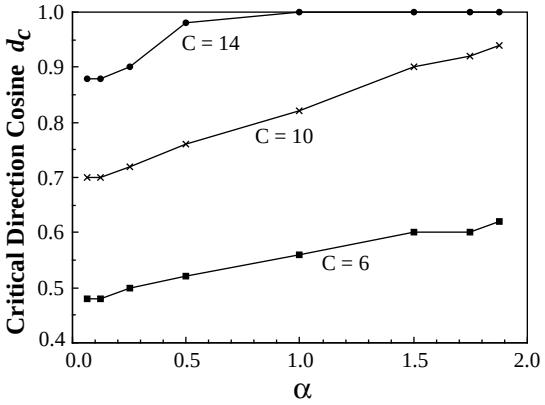

Figure 2. The critical direction cosines d_c for different α values in the DB-PANM ($L = 5$, $\bar{\kappa} \approx 0$, $\bar{\sigma} = 0.4$, and $\alpha = 0.0625, 0.125, 0.25, 0.5, 1.0, 1.5, 1.75, 1.875$). The smaller d_c means the larger size of basins of attraction on average. The three lines represent the results for the different number of clusters C.

For comparative purposes, the association performances of the two different pseudoinverse-type ANMs are also evaluated. The one is the conventional pseudoinverse-type ANM whose connection matrix is given by $\boldsymbol{X}\boldsymbol{X}^{+}$. For the notational convenience, this ANM is denoted by PANM-1. The other is the Gorodnicy's model (denoted by PANM-2) in which the desaturating coefficient D is set to 0.1 in Equation (11). Although the above two ANMs are originally proposed as discrete-time models, we apply continuous-time dynamics to these models in order to compare them with the DB-PANM.

The parameters for the Type-I pattern sets are as follows: $C = 2 \cdots 16$, $L = 5$, $\bar{\kappa} \approx 0$, and $\bar{\sigma} = 0.4$ or 0.7. The α value for the DB-PANM is set to 0.125 (this is an optimal value obtained from the previous simulation results). Figures 3ab show the association performances of the three ANMs for the Type-I pattern sets when σ values are respectively 0.4 and 0.7. For all the pattern sets, the DB-PANM realizes broader basins of attraction as compared with the others as long as C is not too large. Figure 4 demonstrates the association performance for the alphabet patterns. The horizontal axis means the initial direction cosines d that is calculated by Equation (14), and the vertical axis means the success rate of recollections. In this experiment, one can say that the DB-PANM also have broader basins of attraction.

We also should notice that the PANM-2 greatly outperforms the PANM-1 in both of the experiments. Since the diagonal elements of the connection matrix are reduced in the PANM-2, all the eigenvalues are smaller than those of the PANM-1 [6]; as a result, the PANM-2

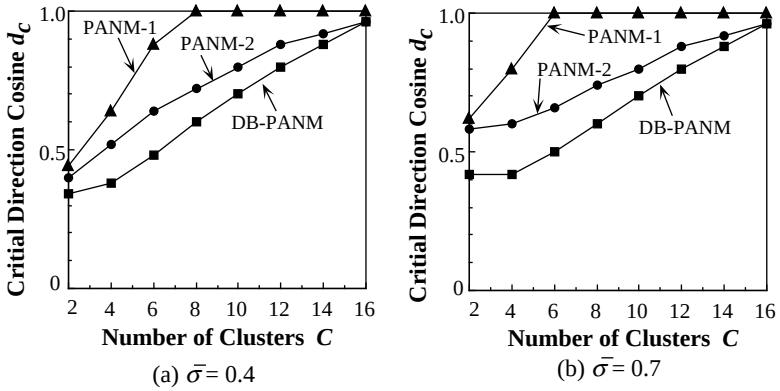

Figure 3. The critical direction cosines d_c for the number of clusters C when the average correlation $\bar{\sigma}$ is set to (a)0.4 or (b)0.7 ($L = 5$, $\bar{\kappa} \approx 0$). The three lines represent the results for PANM-1, PANM-2 and DB-PANM.

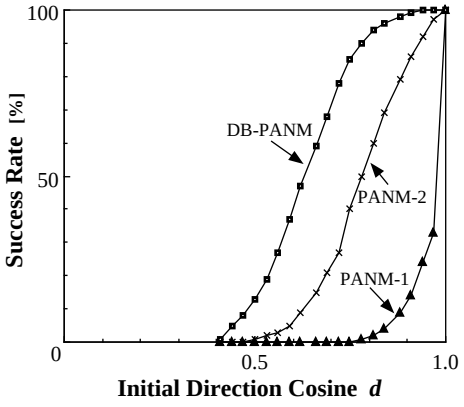

Figure 4. The association performances of PANM-1, PANM-2 and DB-PANM for 26 alphabet patterns. Note that the performances are represented as success rate instead of critical direction cosine.

tends to have negative eigenvalues in the noise-subspace (note that the noise-subspace of the PANM-2 is generally different from that of the DB-PANM). From the results for the PANM-2 and the DB-PANM, it is confirmed that the reduction of the noise-subspace components is the key to improve the performance in the continuous-time model of ANMs.

4. Conclusions

We presented a continuous-time model of ANMs that corresponded to a modified version of the conventional pseudoinverse-type ANM. This model (denoted by DB-PANM) was derived from minimizing the energy function for a modular neural network. Through the eigendecomposition of the connection matrix, we showed that the dynamics of the DB-PANM was classified into two types of the subspace dynamics: the pattern-subspace dynamics and the noise-subspace dynamics. The pattern-subspace is defined as a subspace spanned by memory patterns, while the noise-subspace is defined as an orthogonal complement of the pattern-subspace. The state variation in the pattern-subspace is featured by positive eigenvalues of the connection matrix, hence the pattern-subspace components of network states are monotonously augmented by the network dynamics. On the contrary, the state variation in the noise-subspace is characterized by the negative eigenvalues, thus all the noise-subspace components are monotonously reduced. Since the noise-subspace components are unnecessary for correct recollections, one can say that the noise-subspace dynamics in the DB-PANM realizes the noise reduction for corrupted inputs. The positive eigenvalues for the pattern-subspace can be changed by the parameter α, therefore the dynamical nature of the DB-PANM is also variable according to this parameter.

In the first simulation, we investigated the influence of α to the association performance. As a result, we found that α had an optimal value which gave the largest size of basins of attraction. Furthermore, the basins tend to be smaller as α becomes larger. This fact suggests that it is possible to control the size of attractive basins to some extent by changing α. In the second simulation, we investigated the performance robustness of the DB-PANM for two types of correlated pattern sets: randomly generated cluster patterns and 26 alphabet patterns. When we adopted an optimal value of α, the DB-PANM realized the high association performance as compared with the two versions of pseudoinverse-type ANMs. The simulation results confirm the usefulness of the proposed ANM.

References

1. J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Aca. Sci. U.S.A., Vol. 79, pp.2254–2558, 1982.
2. T. Kohonen, Self-organization and associative memory (3rd ed.), Springer Verlag Berlin, 1989.

3. S. Amari, "Neural theory of association and concept-formation", Biol. Cybern., Vol.26, pp.175–185, 1977.
4. M.H. Hassoun and A.M. Youssef, "High performance recording algorithm for Hopfield model associative memories", Optical Engineering, Vol.28, pp.46–54, 1989.
5. B. Telfer and D. Casasent, "Ho-Kashyap optical associative processors", Applied Optics, Vol.29, pp.1191–1202, 1990.
6. I. Kanter and H. Sompoilnsky, "Associative recall of memory without errors", Physical Review A, Vol.35, pp.380–392, 1987.
7. Y-C. Ho & R.L. Kashyap, "An algorithm for linear inequalities and its applications", IEEE Trans. on Electronic Computers, Vol.EC-14, pp.683–688, 1965.
8. K. Tsutsumi, "Cross-Coupled Hopfield Nets via generalized-delta-rule-based internetworks", Proc. of Int. Joint Conf. on Neural Networks (IJCNN90-San Diego), Vol.II, pp.259–265, 1990.
9. D.O. Gorodnichy and A.M. Reznik, "Increasing attraction of pseudo-inverse autoassociative networks", Neural Processing Letters, Vol.5, pp.121-125, 1997.
10. K. Tsutsumi, "Higher degree error backpropagation in Cross-Coupled Hopfield Nets", Proc. of Int. Joint Conf. on Neural Networks (IJCNN91-Seattle), Vol.II, pp.349–355, 1991.
11. S. Ozawa, K. Tsutsumi, and N. Baba, "An artificial modular neural network and its basic dynamical characteristics", Biol. Cybern., Vol.78, pp.19–36, 1998.
12. K. Ikeda, "A spurious-memory free associative memory system – hysteresis neurons and pseudo-inverse matrix model", The Brain & Neural Networks (in Japanese), Vol.3, pp.141–146, 1996.
13. D.O. Gorodnichy, "A way to improve error correction capability of Hopfield associative memory in the case of saturation", HELNET International Workshop on Neural Networks Proceedings (HELNET 94-95), Vol.I/II, pp.198-212, 1996.
14. T. Kindo and H Kakeya, "A geometrical analysis of associative memory", Neural Networks, Vol.11, pp.39–51, 1998.
15. S. Ozawa, K. Tsutsumi, and N. Baba, "Association performance of Cross-Coupled Hopfield Nets for correlated patterns", Proc. of Int. Joint Conf. on Neural Networks (IJCNN93-Nagoya), Vol.II, pp.2335–2338, 1993.

Appendix

A. Derivation of Network Dynamics

Let us define the row vectors $\boldsymbol{W}_i$ and $\boldsymbol{I}_i$ $(i = 1, \cdots, N)$ represented by the following relations:

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2 \\ \vdots \\ \boldsymbol{W}_N \end{pmatrix} \qquad \boldsymbol{I} = \begin{pmatrix} \boldsymbol{I}_1 \\ \boldsymbol{I}_2 \\ \vdots \\ \boldsymbol{I}_N \end{pmatrix}.$$

Then, Equation (2) is rewritten as follows:

$$\frac{dE}{dt} = -\frac{d\boldsymbol{v}'}{dt}\left[\alpha \boldsymbol{T}\boldsymbol{v} - (\boldsymbol{I} - \boldsymbol{W})'(\boldsymbol{I} - \boldsymbol{W})\boldsymbol{v}\right]$$

$$
\begin{aligned}
&+\beta \cdot \mathrm{tr}\left[\begin{pmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2 \\ \vdots \\ \boldsymbol{W}_N \end{pmatrix} \begin{pmatrix} \frac{d\boldsymbol{W}'_1}{dt} & \frac{d\boldsymbol{W}'_2}{dt} & \cdots & \frac{d\boldsymbol{W}'_N}{dt} \end{pmatrix}\right] \\
&-\boldsymbol{v}' \begin{pmatrix} \frac{d\boldsymbol{W}'_1}{dt} & \frac{d\boldsymbol{W}'_2}{dt} & \cdots & \frac{d\boldsymbol{W}'_N}{dt} \end{pmatrix} \begin{pmatrix} (\boldsymbol{I}_1 - \boldsymbol{W}_1) \\ (\boldsymbol{I}_2 - \boldsymbol{W}_2) \\ \vdots \\ (\boldsymbol{I}_N - \boldsymbol{W}_N) \end{pmatrix} \boldsymbol{v} \\
= & -\frac{d\boldsymbol{v}'}{dt}\left[\alpha \boldsymbol{T}\boldsymbol{v} - (\boldsymbol{I}-\boldsymbol{W})'(\boldsymbol{I}-\boldsymbol{W})\boldsymbol{v}\right] \\
&+\sum_{i=1}^{N}\left[-\boldsymbol{v}'\frac{d\boldsymbol{W}'_i}{dt}(\boldsymbol{I}_i - \boldsymbol{W}_i)\boldsymbol{v} + \beta \boldsymbol{W}_i \frac{d\boldsymbol{W}'_i}{dt}\right] \\
= & -\frac{d\boldsymbol{v}'}{dt}\left[\alpha \boldsymbol{T}\boldsymbol{v} - (\boldsymbol{I}-\boldsymbol{W})'(\boldsymbol{I}-\boldsymbol{W})\boldsymbol{v}\right] \\
&+\sum_{i=1}^{N}\frac{d\boldsymbol{W}_i}{dt}\left[-\boldsymbol{v}(\boldsymbol{I}_i - \boldsymbol{W}_i)\boldsymbol{v} + \beta \boldsymbol{W}'_i\right]. \qquad (15)
\end{aligned}
$$

Note that $\boldsymbol{W}_i$ is a row vector and $\boldsymbol{v}$ is a column vector. In order to ensure the inequality $d\boldsymbol{E}/dt \leq 0$, it is sufficient that we select the network dynamics as follows:

$$c\frac{d\boldsymbol{u}}{dt} = \{\alpha \boldsymbol{T} - (\boldsymbol{I}-\boldsymbol{W})'(\boldsymbol{I}-\boldsymbol{W})\}\boldsymbol{v} \qquad (16)$$

$$\eta\frac{d\boldsymbol{W}_i}{dt} = -\left[\boldsymbol{W}'_i - \frac{1}{\beta}\boldsymbol{v}(\boldsymbol{I}-\boldsymbol{W}_i)\boldsymbol{v}\right]' = -\boldsymbol{W}_i + \frac{1}{\beta}(\boldsymbol{I}-\boldsymbol{W}_i)\boldsymbol{v}\boldsymbol{v}'. \qquad (17)$$