Efficient Continuous-Time Asymmetric Hopfield Networks for Memory Retrieval

Pengsheng Zheng  
zhengpengsheng@yahoo.cn  
Wansheng Tang  
tang@tju.edu.cn  
Jianxiong Zhang  
jxzhang@tju.edu.cn  
Institute of Systems Engineering, Tianjin University, 300072, China

A novel $m$ energy functions method is adopted to analyze the retrieval property of continuous-time asymmetric Hopfield neural networks. Sufficient conditions for the local and global asymptotic stability of the network are proposed. Moreover, an efficient systematic procedure for designing asymmetric networks is proposed, and a given set of states can be assigned as locally asymptotically stable equilibrium points. Simulation examples show that the asymmetric network can act as an efficient associative memory, and it is almost free from spurious memory problem.

1 Introduction

Memory retrieval in artificial neural networks has been extensively investigated in the past decades, and many interesting results have been obtained. For example, Amari (1972a) studied the multistable and oscillatory behaviors of random nets with analog neuron-like elements. Both symmetric and asymmetric networks are studied in Amari (1972b). Primitive neural models of concept formation and characteristics of the learning rules are studied in Amari (1977). Strange dynamical behaviors in recalling processes have been investigated in Amari and Maginu (1988). Hopfield (1982, 1984) introduced the concept of an energy function for studying the stability of the network and found that the recurrent artificial neural network appears to be a powerful tool for associative memory. The conventional networks with symmetric connections were intensively studied. However, the synaptic couplings are in general asymmetric in biological nervous systems, and Hertz, Grinstein, and Solla (1987) found that it was almost impossible to implement the hardware network precisely with symmetric connections, so studies of asymmetric networks are of particular biological and practical importance.

The dynamical behavior analysis, especially the retrieval property analysis, of the asymmetric Hopfield networks is a primary step for the practical
application of the network. Amit, Gutfreund, and Sompolinsky (1985) proposed that small asymmetry in the synaptic connections did not qualitatively change the behavior of the system. Fukai and Shiino (1990) studied noise-driven chaotic motions in the retrieval process. Sompolinsky and Kanter (1986) found that asymmetric neural networks could provide temporal association memory. Cheng, Dasgupta, and Singh (2000) studied the retrieval properties of Hopfield model with random asymmetric interactions by computer simulations. Pattern recognition using asymmetric discrete-time attractor neural networks was introduced in Jin and Zhao (2005). Xu, Hu, and Kwong (1996) obtained some sufficient conditions for the discrete-time asymmetric Hopfield networks converging to a stable state. Some sufficient conditions for the global and local stability of the network have been proposed in Michel and Gray (1990), Yang and Dillon (1994), Matsuoka (1992), Chen and Amari (2001), and Liu and Chen (2002). However, most of these results are too complicated for application. In this letter, we propose two concise results by using an explicit energy function method. The network dynamic is simply expressed as a downhill motion in the energy landscapes, which can bring us a better understanding to the network dynamic.

In addition, the symmetric Hopfield network can be easily defined by using Hebb’s rule (Hopfield, 1982; Hebb, 1949) or the pseudo-inverse rule (Personnaz, Guyon, & Drefus, 1985). Dealing with asymmetric networks, Zhao (2004) introduced a strategy for designing weights of asymmetric networks with controllable degree of symmetry. Strategies with the concept of higher-order Hamming stability have been proposed in Lee and Chuang (2005) and Zhuang, Huang, and Yu (1994). Learning rules with optimal stability were studied in Krauth and Mezard (1987). Bornholdt and Graudenz (1992) designed general asymmetric networks by using a genetic algorithm. However, they all focused on Hopfield networks with discrete-time dynamics. It is still hard to construct a continuous-time asymmetric network lacking an efficient synthesis procedure.

Therefore, in this letter, the retrieval property of the continuous-time asymmetric Hopfield network is analyzed by using a novel $m$ energy functions method. Sufficient conditions for the local and global asymptotic stability of the network are proposed. Moreover, to make the asymmetric Hopfield networks more applicable, an efficient, systematic system-designing procedure is proposed by using Schur decomposition and single value decomposition, which allows one to assign some predetermined states to the locally asymptotically stable equilibrium points of the asymmetric network. Moreover, the retrieval reliability, spurious memory, and retrieval time of the asymmetric network are also discussed by numerical simulations. We show that the designed network has a perfect retrieval reliability, and it is almost free from spurious memory problem, which outperforms most existing networks.
2 Retrieval Property Analysis

Consider an asymmetric Hopfield network of the form

$$\dot{X} = -AX + W\phi(X) + I,$$

(2.1)

where $X = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$ is the neuron state vector, $I = [I_1, I_2, \ldots, I_n]^T \in \mathbb{R}^n$ is a real constant vector, $A = \text{diag}[a_1, a_2, \ldots, a_n] \in \mathbb{R}^{n \times n}$ is a positive diagonal matrix, $W \in \mathbb{R}^{n \times n}$ is a synaptic weights matrix that is asymmetrically defined, a continuous function $\phi: \mathbb{R}^n \to \mathbb{R}^n$, with $\phi(0) = 0$ and $\phi(X) = [\phi_1(x_1), \phi_2(x_2), \ldots, \phi_n(x_n)]^T$, denotes the neuron transfer function, and $\phi_j(x)$ is differentiably defined and satisfies the following condition,

$$0 < \frac{\phi_j(x) - \phi_j(y)}{x - y} \leq k, \quad \forall x, y \in \mathbb{R}, \quad j = 1, 2, \ldots, n.$$ 

(2.2)

It follows that $\phi_j(\cdot)$ ($j = 1, 2, \ldots, n$) are monotonically increasing, and $\phi_j(x) = 0$ if and only if $x = 0$. In addition, network 2.1 appears to be a linear system if $\phi(\cdot)$ is a linear function. It is impossible to endow a linear system with retrieval properties. Hence, in this letter, we consider $\phi(\cdot)$ to be a nonlinear function.

Let $S = \{X^{(i)}, i = 1, 2, \ldots, m\} \subset \mathbb{R}^n$ be the equilibrium points set of the asymmetric network 2.1. We restrict $m \geq 1$: that is, the network has at least one equilibrium point.

To endow the asymmetric network with retrieval properties, a given set of states needs to be designed as equilibrium points of the network. Moreover, the network is expected to converge to an equilibrium point if the initial state is sufficiently close to it (i.e., the equilibrium points are locally asymptotically stable). In the following, a sufficient condition for the local asymptotical stability of the equilibrium points is proposed.

**Theorem 1.** $X^{(i)}$ is locally asymptotically stable if

$$H - A\Phi(X^{(i)}) < 0,$$

(2.3)

where $\Phi(X) = \text{diag}[1/\phi_1(x_1), \ldots, 1/\phi_n(x_n)]$ and $H = (W + W^T)/2$.

**Proof.** Consider $m$ continuously differentiable energy functions as

$$V^{(i)}(U) = \sum_{j=1}^n \int_0^{u_j} g_j(s) \, ds, \quad i = 1, 2, \ldots, m,$$

(2.4)
where \( U = X - X^{(i)} = [u_1, u_2, \ldots, u_n]^T \), and

\[
g_j(u_j) = \varphi_j(u_j + x_j^{(i)}) - \varphi_j(x_j^{(i)}), \quad j = 1, 2, \ldots, n. \tag{2.5}
\]

Taking one energy function, \( V^{(i)}(U) \), for example, the following analysis holds for \( i = 1, 2, \ldots, m \).

Asymmetric network 2.1 is transformed into

\[
\dot{U} = -AU + Wg(U). \tag{2.6}
\]

Let \( U^{(i)} \) be the stationary point of \( V^{(i)}(U) \). Thus,

\[
\frac{\partial V^{(i)}(U)}{\partial u_j} \bigg|_{u_j = u_j^{(i)}} = g_j(u_j^{(i)}) = 0, \quad j = 1, 2, \ldots, n.
\]

It follows from equations 2.2 and 2.5 that \( g_j(\cdot) \ (j = 1, 2, \ldots, n) \) are nonlinear differentiable functions. Moreover, \( g_j(\cdot) \ (j = 1, 2, \ldots, n) \) increase monotonically, and \( g_j(u) = 0 \) if and only if \( u = 0 \). Hence, \( U^{(i)} = 0 \) is the unique stationary point of \( V^{(i)}(U) \) and

\[
\frac{\partial^2 V^{(i)}(U)}{\partial U^2} = \text{diag}[\dot{g}_1(u_1), \ldots, \dot{g}_n(u_n)] \geq 0
\]

for any \( U \in \mathbb{R}^n \). In addition, it follows from equation 2.4 that \( V^{(i)}(U^{(i)}) = V^{(i)}(0) = 0 \). Therefore, \( V^{(i)}(U) \geq 0 \) and \( V^{(i)}(U) = 0 \) if and only if \( U = U^{(i)} = 0 \). This implies that \( U^{(i)} = 0 \) is the global minimum point of \( V^{(i)}(U) \).

Moreover, if \( U = U^{(i)} = 0 \), then \( X = X^{(i)} \). This implies that the global minimum point of \( V^{(i)}(U) \) corresponds to the equilibrium point \( X^{(i)} \) of network 2.1.

It follows from equations 2.2 and 2.5 that

\[
\dot{g}_j(u_j^{(i)}) = \dot{g}_j(0) = \dot{\varphi}_j(x_j^{(i)}), \quad j = 1, 2, \ldots, n.
\]

The linearization of a nonlinear system at the equilibrium point is a powerful tool for local stability analysis. Define \( B_{\varepsilon_i}(U^{(i)}) = \{ U : \| U - U^{(i)} \| < \varepsilon_i \} \). If \( \varepsilon_i \) is sufficiently small, then \( g(U) \) can be represented by its linearization at \( U^{(i)} \), such that

\[
U = \Phi(X^{(i)})g(U). \tag{2.7}
\]

Let \( H = (W + W^T)/2 \), and \( H, W \in \mathbb{R}^{n \times n} \). For any \( X \in \mathbb{R}^n \), it holds that

\[
X^T WX + (X^T WX)^T = 2X^T WX,
\]

\[
X^T WX + X^T W^TX = 2X^T HX;
\]
Asymmetric Hopfield Networks for Memory Retrieval

Thus,
\[ X^T H X = X^T W X. \] (2.8)

When \( \epsilon_i \) is sufficiently small and \( U \in B_{\epsilon_i}(U^{(i)}) \), it follows from equations 2.4 and 2.6 to 2.8 that the time derivative of \( V^{(i)}(U) \) along the trajectories of the network can be described as
\[
\dot{V}^{(i)}(U) = g^T(U) \dot{U} \\
= g^T(U)[-AU + Wg(U)] \\
= -g^T(U)AU + g^T(U)Wg(U) \\
= -g^T(U)A\Phi(X^{(i)})g(U) + g^T(U)Hg(U) \\
= g^T(U)(H - A\Phi(X^{(i)}))g(U). \] (2.9)

Hence, in a certain neighborhood of \( U^{(i)} \), if \( H - A\Phi(X^{(i)}) < 0 \), then \( \dot{V}^{(i)}(U) \leq 0 \), and \( \dot{V}^{(i)}(U) = 0 \) if and only if \( g(U) = 0 \), that is, \( U = U^{(i)} = 0 \). Then, given an initial state \( U(0) \) that is sufficiently close to \( U^{(i)} \), as the network dynamic evolves with time, \( V^{(i)}(U) \) keeps decreasing until it reaches its global minimum point \( U^{(i)} \), which also corresponds to the equilibrium point \( X^{(i)} \) of network 2.1. According to the Lyapunov stability criterion, \( X^{(i)} \) is locally asymptotically stable.

Network 2.1 can be endowed with a retrieval property if \( X^{(i)} (i = 1, 2, \ldots, m) \) are designed as locally asymptotically stable equilibrium points. It should be noted that local stability concerns only the behavior in a certain neighborhood of the equilibrium point. In the following, we propose a sufficient condition for the global asymptotic stability of network 2.1:

**Theorem 2.** If network 2.1 has an equilibrium point and
\[
H - A/k < 0, \tag{2.10}
\]
\( H = (W + W^T)/2. \) Then network 2.1 is globally asymptotically stable, and it has a unique equilibrium point.

**Proof.** Supposing network 2.1 has \( m \) equilibrium points, we can define \( m \) energy functions as shown in equation 2.4. Taking one energy function \( V^{(i)}(U) \), for example, the following analysis holds for \( i = 1, 2, \ldots, m \).

Define \( Q_k = \text{diag}[g_1(u_1)/u_1, \ldots, g_n(u_n)/u_n] \) with \( g_j(0)/0 = \dot{g}_j(0) \) and \( P_k = \text{diag}[k, \ldots, k] \in \mathbb{R}^{n \times n} \).

It follows from equations 2.2 and 2.5 that
\[
0 < \frac{g_j(u_j)}{u_j} \leq k, \quad j = 1, 2, \ldots, n \tag{2.11}
\]
and
\[
g_j(u)u \geq 0, \quad \forall u \in \mathbb{R}, \quad j = 1, 2, \ldots, n. \tag{2.12}
\]

It follows from equation 2.11 that
\[
P_k - Q_k \geq 0. \tag{2.13}
\]

From the fact that \( A \) is a positive diagonal matrix, it follows from equations 2.8, 2.9, 2.12, and 2.13 that
\[
\dot{V}^{(i)}(U) = -g^T(U)\frac{A}{k}P_k U + g^T(U)Wg(U)
\leq -g^T(U)\frac{A}{k}Q_k U + g^T(U)Wg(U)
= -g^T(U)\frac{A}{k}g(U) + g^T(U)Hg(U)
= g^T(U)\left(H - \frac{A}{k}\right)g(U).
\]

Thus, if \( H - A/k < 0 \), then \( \dot{V}^{(i)}(U) \leq 0 \), and \( \dot{V}^{(i)}(U) = 0 \) if and only if \( g(U) = 0 \), that is, \( U = U^{(i)} = 0 \). This implies, given an arbitrary initial state, that all the energy functions \( V^{(i)}(U) \) \((i = 1, 2, \ldots, m)\) decrease monotonically along the solution of the network; as \( t \to +\infty \), all the \( V^{(i)}(U) \) \((i = 1, 2, \ldots, m)\) eventually reach its global minimum value at \( U^{(i)} \). Then it holds that
\[
X^{(1)} = \cdots = X^{(m)} = \lim_{t \to +\infty} X(t).
\]

Hence, network 2.1 has a unique equilibrium point. According to the Lyapunov stability criterion, network 2.1 is globally asymptotically stable.

**Remark 1.** It is worth mentioning that some significant results have been proposed in Matsuoka (1992), Chen and Amari (2001), and Liu and Chen (2002). However, theorems 1 and 2 appear to be more concise and applicable. Moreover, if \( a_1 = \cdots = a_n = 1 \) and \( k = 1 \), theorem 2 reduces to the result in Matsuoka (1992).

**Remark 2.** Theorems 1 and 2 also hold when \( W \) is symmetrically defined, where \( H = W \). It is not hard to verify that equations 2.3 and 2.10 are sufficient conditions for proving the energy function (see equation 7 in Hopfield, 1984) has local minima and global minimum at the equilibrium point, respectively (see appendix A).
3 System Designing

One of the most important applications of Hopfield networks is for recognition of binary patterns. Let \( S = \{ X^{(i)}, i = 1, 2, \ldots, m \} \subset \mathbb{R}^n \) be the vector set that needs to be stored. To endow the asymmetric network with retrieval properties, one should properly assign the parameters \( A, W, \text{ and } I \) such that \( X^{(i)} (i = 1, 2, \ldots, m) \) can be assigned as locally asymptotically stable equilibrium points. For simplicity, we make the following three assumptions:

\[
a_1 = \cdots = a_n \geq 0. \tag{3.1}
\]

\[
\varphi_1(x) = \cdots = \varphi_n(x), \forall x \in \mathbb{R}. \tag{3.2}
\]

\[
| x^{(i_1)}_{j_1} | = | x^{(i_2)}_{j_2} |, \quad i_1, i_2 = 1, 2, \ldots, m, \quad j_1, j_2 = 1, 2, \ldots, n. \tag{3.3}
\]

Let \( \zeta^{(i)} = \varphi(X^{(i)}) \) and \( \Theta = [\zeta^{(1)} - \zeta^{(m)}, \zeta^{(2)} - \zeta^{(m)}, \ldots, \zeta^{(m-1)} - \zeta^{(m)}] \). Performing single value decomposition of \( \Theta \) gives \( \Theta = V \Sigma Z^T \), with \( V \) and \( Z \) being orthogonal matrices and \( \Sigma \in \mathbb{R}^{n \times (m-1)} \) being a diagonal matrix. Then we have the following result:

**Proposition 1.** For \( i = 1, 2, \ldots, m \), we have

\[
W \zeta^{(i)} + I = c \zeta^{(i)},
\]

where \( c \) is a real constant if

\[
W = c VL V^T, \tag{3.4}
\]

\[
I = c \zeta^{(m)} - W \zeta^{(m)}, \tag{3.5}
\]

where \( L = [l_{ij}] \in \mathbb{R}^{n \times n} \) is a lower triangular matrix, with

\[
l_{ij} = \begin{cases} 
1, & \text{for } i = j, i \leq m - 1 \\
\text{rand}(\beta_{ij}, \gamma_{ij}), & \text{for } i > j, i \geq m, j \geq m \\
\tau_i, & \text{for } i = j, i > m - 1 \\
0, & \text{otherwise},
\end{cases}
\]

where \( \beta_{ij}, \gamma_{ij}, \tau_i \in \mathbb{R}, \text{ and } \text{rand}(\beta_{ij}, \gamma_{ij}) \) denote random values within interval \((\beta_{ij}, \gamma_{ij})\).

**Proof.** It is not hard to verify that

\[
W \Theta = c VL V^T V \Sigma Z^T = c VL \Sigma Z^T = c V \Sigma Z^T = c \Theta.
\]
It follows that

\[ W\zeta^{(i)} + I = W(\zeta^{(i)} - \zeta^{(m)} + \zeta^{(m)}) + I \]
\[ = W(\zeta^{(i)} - \zeta^{(m)}) + W\zeta^{(m)} + I \]
\[ = c[\zeta^{(i)} - \zeta^{(m)}] + W\zeta^{(m)} + c\zeta^{(m)} - W\zeta^{(m)} \]
\[ = c\zeta^{(i)}. \]

for \( i = 1, 2, \ldots, m \).

It should be noted that \( W = cVLV^T = VcLV^T \) can be considered as a Schur decomposition of \( W \); therefore, \( c \times l_{jj} \) (\( j = 1, 2, \ldots, n \)) are the eigenvalues of \( W \). Hence, we can easily control the eigenvalues of \( W \) by adjusting parameters \( c \) and \( l_{jj} \).

The system-designing procedure can be concluded as follows

**Step 1:** Compute \( \Theta = [\zeta^{(1)} - \zeta^{(m)}, \ldots, \zeta^{(m-1)} - \zeta^{(m)}] \) where \( \zeta^{(i)} = \varphi(X^{(i)}) \).

**Step 2:** Perform a single value decomposition of \( \Theta: \Theta = V\Sigma Z^T \).

**Step 3:** Set lower triangular matrix \( L \) with \( \tau_i \leq 0 \).

**Step 4:** Compute \( c = a_1x_1^{(1)}/\varphi_1(x_1^{(1)}) \).

**Step 5:** Compute \( W = cVLV^T \), and \( I = c\zeta^{(m)} - W\zeta^{(m)} \).

It follows from proposition 1 that

\[ W\varphi(X^{(i)}) + I = c\varphi(X^{(i)}), \quad i = 1, 2, \ldots, m. \]

Moreover, it follows from equations 3.1 to 3.3 that

\[ c = a_i x_j^{(i)}/\varphi_j(x_j^{(i)}), \quad \forall 1 \leq i \leq m, \quad 1 \leq j \leq n. \]

Hence,

\[ W\varphi(X^{(i)}) + I = a_i x_j^{(i)}/\varphi_j(x_j^{(i)})\varphi(X^{(i)}) = AX^{(i)}. \]

Thus, \( X^{(i)} \) (\( i = 1, 2, \ldots, m \)) can be designed as equilibrium points of the network. To minimize the maximum eigenvalue of \( W \), we set \( \tau_i \leq 0 \). Then \( \lambda_1(W) = c \), where \( \lambda_1(\cdot) \) denotes the maximum eigenvalue of a matrix.

**Remark 3.** The asymmetric networks with \( I = [0, 0, \ldots, 0]^T \) can be constructed by performing the designing procedure from step 2 with \( \Theta = [\zeta^{(1)}, \zeta^{(2)}, \ldots, \zeta^{(m)}] \). Moreover, a symmetric network can be designed by setting \( l_{ij} = 0 \) (\( i > j, i \geq m, j \geq m \)).
4 Simulations

4.1 Example 1. We illustrate the effectiveness of theorem 2 in this example.
Consider an asymmetric network 2.1 with $n = 3$, $a_1 = a_2 = a_3 = 1$, $I_1 = I_2 = I_3 = 0$, $\varphi_1(\cdot) = \varphi_2(\cdot) = \varphi_3(\cdot) = f(\cdot)$ with

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

(4.1)

and $W = \begin{bmatrix} -2 & 1.5 & 2 \\ 1 & -1 & -2 \\ -1.3 & 0.5 & -2 \end{bmatrix}$.

Hence, we have $a_i/k = 1 \ (i = 1, 2, 3)$, and

$$H = \begin{bmatrix} -2 & 1.25 & 0.35 \\ 1.25 & -1 & -0.75 \\ 0.35 & -0.75 & -2 \end{bmatrix}.$$ 

It is not hard to verify that $H - A/k < 0$, which satisfies condition 2.10. $X^{(1)} = 0$ is an equilibrium point of the network. According to theorem 2, $X^{(1)} = 0$ is the unique equilibrium point of the defined network, and the network is globally asymptotically stable. Furthermore, 100 initial states, which are uniformly distributed within $B_0 = \{X : -6 \leq x_i \leq 6, \ i = 1, 2, 3\}$, are presented to the network. As the network dynamic evolves with time, they all finally converge to $X^{(1)} = 0$. This illustrates the effectiveness of theorem 2.

4.2 Example 2. We illustrate the use of the system-designing procedure in this example.
The four patterns needing to be stored are specified by

$$X^{(1)} = [-0.8, 0.8, 0.8, 0.8, 0.8, -0.8, -0.8, 0.8]^T$$

$$X^{(2)} = [0.8, -0.8, 0.8, -0.8, 0.8, 0.8, -0.8, -0.8]^T$$

$$X^{(3)} = [0.8, 0.8, -0.8, -0.8, 0.8, -0.8, -0.8, -0.8]^T$$

$$X^{(4)} = [-0.8, -0.8, 0.8, -0.8, -0.8, 0.8, 0.8, -0.8]^T$$

Then $S = \{X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)}\}$. We set $a_1 = \cdots = a_8 = 1$, and $\varphi_1(\cdot) = \cdots = \varphi_8(\cdot) = f(\cdot)$. Then a proper network needs to be designed with $W\varphi(X^{(i)}) + I = X^{(i)}, \ i = 1, 2, 3, 4.$
It follows from the system-designing procedure that

\[
c = a_1 x_1^{(1)} / \varphi_1(x_1^{(1)}) = 1.205,
\]

\[
\Theta = \begin{bmatrix}
0 & 1.33 & 1.33 \\
1.33 & 0 & 1.33 \\
0 & 0 & -1.33 \\
1.33 & 0 & 0 \\
1.33 & 1.33 & 1.33 \\
-1.33 & 0 & 0 \\
-1.33 & 0 & -1.33 \\
1.33 & 0 & 0
\end{bmatrix}.
\]

We set \( \tau_i = 0 \) (3 < \( i \) ≤ 8), and \( \beta_{ij} = -0.5, \gamma_{ij} = 0.5, \lambda_{ij}(i > j, i \geq 4, j \geq 4) \) are randomly produced within interval \((-0.5, 0.5)\). Then \( W \) and \( I \) can be calculated by equations 3.4 and 3.5.

\[
W = \begin{bmatrix}
0.646 & -0.037 & -0.072 & -0.160 & 0.559 & 0.182 & 0.035 & -0.144 \\
-0.294 & 0.462 & -0.188 & -0.033 & 0.294 & 0.143 & -0.554 & 0.070 \\
-0.290 & -0.357 & 0.665 & -0.287 & 0.290 & -0.373 & 0.183 & 0.164 \\
-0.273 & 0.103 & 0.277 & 0.130 & 0.273 & -0.251 & -0.173 & 0.275 \\
0.559 & 0.037 & 0.072 & 0.160 & 0.646 & -0.182 & -0.035 & 0.144 \\
0.133 & 0.083 & -0.103 & -0.558 & -0.133 & 0.193 & 0.185 & -0.217 \\
-0.004 & -0.386 & 0.351 & 0.254 & 0.004 & 0.516 & 0.467 & -0.094 \\
0.137 & 0.299 & 0.088 & 0.643 & -0.137 & -0.206 & 0.210 & 0.405
\end{bmatrix},
\]

\[
I = [-0.32, -0.07, 0.17, -0.36, 0.32, 0.16, -0.24, 0.03]^T.
\]

It is not hard to verify that \( \lambda_i(H - A\Phi(X^{(i)})) = -0.58 < 0 \) for \( i = 1, 2, 3, 4 \), which satisfies the condition of theorem 1. Hence \( X^{(i)}(i = 1, 2, 3, 4) \) can be designed as locally asymptotically stable equilibrium points.

Given an initial state \( X_1(0) = [7.0, -6.6, 8.5, -5.6, 8.0, -6.5, 7.8, -9.3]^T \), the network dynamic evolves with time, and the network finally converges to \( X^{(2)} \). Figure 1 illustrates the trajectories of the network.

The spurious memories may exist in the designed network, which do harm to the network performance. In the following, we study the spurious memories by numerical simulations.

Assume \( X_1^1 = [0, -0.8, 0.8, -0.8, 0, 0.8, 0.8, -0.8]^T \). It is not hard to verify that \( X_1^1 \) is an equilibrium point of the designed network. With the initial state \( X_2(0) = [0.3, -0.6, 1, -0.7, -0.3, 0.7, 0.7, -0.9]^T \), Figure 2 shows that the network finally converges to \( X_1^1 \), which appears to be a spurious memory.
It is difficult to compute every spurious memory analytically since network 2.1 is nonlinearly coupled. However, the network can converge to a spurious memory if an initial state is sufficiently close to it. Moreover, it can be concluded that the equilibrium points exist only in a certain neighborhood of the origin (see appendix B). Hence, in this example, we try to find all the spurious memories by giving the network different initial states: 7140 vectors, which are uniformly distributed within \( B_s = \{ X : -4 \leq x_i \leq 4, i = 1, 2, \ldots, 8 \} \), are presented to the designed network as initial states. It follows from the results of appendix B that all the equilibrium points (including memorized patterns and spurious memories) are inside \( B_s \). However, only three spurious memories are found:

\[
X^1_s = [0, -0.8, 0.8, -0.8, 0, 0.8, 0.8, -0.8]^T \\
X^2_s = [0, 0.8, -0.8, -0.8, 0, 0.8, -0.8, -0.8]^T \\
X^3_s = [-0.8, 0.8, -0.8, -0.8, -0.8, 0.8, -0.8, -0.8]^T.
\]

It is not hard to verify that

\[
X^1_s = (X^{(2)} + X^{(4)})/2
\]
Figure 2: Phase portrait of the network in example 2 with initial state $X_2(0)$.

$$X^2_s = -X^{(2)}/2 + X^{(3)} + X^{(4)}/2$$

and

$$X^3_s = -X^{(2)} + X^{(3)} + X^{(4)}.$$
Asymmetric Hopfield Networks for Memory Retrieval

Figure 3: Phase portrait of the network in example 2 with initial state $X_3(0)$.

Figure 4: Pattern correspondence to six equilibrium points of the network.

the Jacobian matrix evaluated at $X_u$ have positive real parts. Thus, $X_u$ is not a spurious memory because it is unstable.

4.3 Example 3. We demonstrate the application of asymmetric network 2.1 in this example.

Let us consider patterns made up of $7 \times 8$ small boxes, each pattern corresponding to a 56-dimensional vector, with each component value varying between $-1$ (black) and 1 (white). Figure 4 shows the six patterns that need to be stored as equilibrium points. We perform the designing procedure with $n = 56$, $\varphi_1(\cdot) = \ldots = \varphi_{56}(\cdot) = f(\cdot)$, $c = 1.3$, $a_i = 1$, $\beta_{ij} = -1$, $\gamma_{ij} = 1$, $l_{ij}(i > j, i \geq 6, j \geq 6)$, which are randomly produced within interval $(-1, 1)$, $\tau_i = -3$ ($5 < i \leq 56$). Then $W$ and $I$ can be calculated by equations 3.4 and 3.5.

Six corrupted patterns, as shown in Figure 5a, are presented to the network; during the retrieval phase, the network dynamic evolves toward the corresponding stored patterns. Patterns in Figure 5b illustrate the retrieved
patterns of the corrupted ones in the same column. As shown in Figure 5, the network retrieves a previously stored pattern that most closely resembles the corrupted one. When the simulation results in example 2 are combined, it is easy to see that the designed network has perfect retrieval reliability.

In addition, we assume

$$X_u = \left( X^{(i)} + X^{(j)} \right) / 2, \quad \forall i \neq j, 1 \leq i, j \leq 6,$$

where $X^{(i)} (1 \leq i \leq 6)$ denote the patterns shown in Figure 4.

It can be verified that $X_u$ are equilibrium points of the designed network. However, all are unstable. This implies that $X_u$ are not spurious memories.

Furthermore, 1000 initial states, with random strings of ±1, are presented to the designed network. However, all converge to the stored patterns, and no spurious memory is found.

We select 100 initial states, and the Hamming distance between the initial states and the corresponding remember patterns is 9. In simulations, they all converge to the stored patterns. We also design a discrete time asymmetric network by using the method presented in Lee and Chuang (2005). However, only 31 patterns of 100 initial states can be successfully recalled, and 11 spurious memories are found from the 100 initial states. The network constructed in this letter apparently has a better retrieval reliability.

5 Further Discussion

As indicated in expression 2.9, the retrieval time of a designed network has correlation to the value of $\lambda_1(H - A\Phi(X^{(i)}))$ which is negative. The retrieval time can be reduced by minimizing the value of $\lambda_1(H - A\Phi(X^{(i)}))$. We can achieve this by representing patterns with strings of ±$\omega$ where $\omega > 1$ with $-\omega$(black), $\omega$(white). To illustrate this, we replace the pregiven patterns in example 2 by four new patterns—$2 \cdot X^{(1)}$, $2 \cdot X^{(2)}$, $2 \cdot X^{(3)}$, $2 \cdot X^{(4)}$—and adopt the same system designing parameters as example 2. Then a new asymmetric network can be constructed. It can be calculated that $\Phi(X^{(i)}) = \text{diag}[14.2, \ldots, 14.2], \lambda_1(H - A\Phi(X^{(i)})) = -12.1 < 0$ for $i = 1, \ldots, 6$. $\lambda_1(H - A\Phi(X^{(i)}))$
Asymmetric Hopfield Networks for Memory Retrieval

2, 3, 4, which is much smaller than that of the network designed in example 2. Figure 6 illustrates the trajectories of the network with the initial state \( X(0) = [7.0, -6.6, 8.5, -5.6, 8.0, -6.5, 7.8, -9.3]^T \). When Figure 6 is compared with Figure 1, it is easy to see that the designed network has reduced retrieval time.

In simulations, we find that the designed networks may have isolated nodes if \( X(i) (i = 1, 2, \ldots, m) \) are not well selected, which usually happens if the patterns are linearly dependent. Moreover, condition 3.3 limits the applications of the proposed method to the patterns of bipolar forms. A more general system-designing procedure still needs further research.

As shown in the numerical simulations, the asymmetric network can reach an equilibrium point when an initial state that is sufficiently close to this equilibrium point is given. However, we cannot describe the attraction domain of each equilibrium point precisely. The attraction domain of the asymmetric networks still needs further study.

6 Conclusion

Sufficient conditions for the local and global asymptotic stability of the asymmetric network are proposed by using an explicit energy function...
method, and a systematic system designing procedure is developed by using Schur decomposition. We showed that asymmetric Hopfield-type networks can work as efficient associative memories.

**Appendix A: Proof of Remark 2**

Define the energy function of the symmetric network 2.1 as

\[
V(O) = -\frac{1}{2} OT WO - OT I + \sum_j^n a_j \int_0^{o_j} \varphi_{j}^{-1}(s) ds, \tag{A.1}
\]

where \( O = [o_1, o_2, \ldots, o_n]^T \) with \( o_j = \varphi_i(x_j)(i = 1, 2, \ldots, n) \). Actually, energy function A.1 is identical to equation 2.7 in Hopfield (1984).

Let \( O^{(i)} = \varphi(X^{(i)}) \). For \( i = 1, 2, \ldots, m \), we have

\[
\frac{\partial V}{\partial O^{(i)}_j} = -WO^{(i)}_j + a_j \varphi^{-1}(O^{(i)}_j) - I = 0, \quad j = 1, 2, \ldots, n.
\]

Thus, the stationary points of the energy function A.1 are also equilibrium points of the symmetric network 2.1.

It is not hard to verify that

\[
\varphi_j^{-1}(o_j) = 1/\varphi_j(x_j), \quad j = 1, 2, \ldots, n.
\]

Hence, the second partial derivative of \( V(O) \) can be described as

\[
\frac{\partial^2 V}{\partial O^2} = -W + A \Phi(X).
\]

It is obvious that in a certain neighborhood of \( O^{(i)} \), if \( W - A \Phi(X^{(i)}) < 0 \), then \( -W + A \Phi(X) > 0 \). Hence, \( O^{(i)} \) is a local minimum point of energy function A.1. The analysis holds for \( i = 1, 2, \ldots, m \).

Moreover, if \( W - A/k < 0 \) then \( -W + A \Phi(X) > 0 \) for \( O \in \mathbb{R}^n \). Hence, \( O^{(i)} \) is the global minimum point of equation A.1.

**Appendix B: Domain Estimation of the Equilibrium Points**

The following analysis gives a proof that the equilibrium point exists only in a certain neighborhood of the origin.

Assume \( X^* = [x_1^*, x_2^*, \ldots, x_n^*]^T \) is an equilibrium point of the network, \( B_\rho(X) = \{ X : \| X \| < \rho \} \), and \( B_\rho^C(X) = \mathbb{R}^n \setminus B_\rho(X) \). Then

\[
W \varphi(X^*) + I = AX^*. \tag{B.1}
\]
In example 2, \( a_i = 1 \) and \( \varphi_j(x) = f_1(x) < 1, \forall x \in \mathbb{R} \). Then

\[
\sum_{j=1}^{n} w_{ij} \varphi_j(x_j) + I_i \leq \sum_{j=1}^{n} |w_{ij}| + |I_i|.
\] (B.2)

As shown in equation B.2, if \( \exists 0 \leq i \leq n, \rho = \sum_{j=1}^{n} |w_{ij}| + |I_i| \), then there is no \( X \in B^o_\rho(X) \) satisfying condition B.1. Hence, the equilibria of network 2.1, including memorized patterns and spurious memories, exist only in \( B_\rho(X) \).

References


Received May 9, 2009; accepted October 14, 2009.