Available online at www.sciencedirect.com
SCIENCE DIRECT
Physica D 197 (2004) 134–148
PHYSICA D
www.elsevier.com/locate/physd

Oscillatory associative memory network with perfect retrieval

Takashi Nishikawa[a,1], Frank C. Hoppensteadt[b], Ying-Cheng Lai[a,c,*]

[a] *Department of Mathematics and Center for Systems Science and Engineering Research, Arizona State University, Tempe, Arizona 85287, USA*
[b] *Courant Institute of Mathematical Sciences, New York University, New York NY 10012, USA*
[c] *Department of Electrical Engineering, Arizona State University, Tempe, Arizona 85287, USA*

Received 15 August 2003; received in revised form 2 June 2004; accepted 22 June 2004

Communicated by Y. Kuramoto

Abstract

Inspired by the discovery of possible roles of synchronization of oscillations in the brain, networks of coupled phase oscillators have been proposed before as models of associative memory based on the concept of temporal coding of information. Here we show, however, that error-free retrieval states of such networks turn out to be typically *unstable* regardless of the network size, in contrast to the classical Hopfield model. We propose a remedy for this undesirable property, and provide a systematic study of the improved model. In particular, we show that the error-free capacity of the network is at least $2\varepsilon^2/\log n$ patterns per neuron, where n is the number of oscillators (neurons) and ε the strength of the second-order mode in the coupling function.

PACS: 87.10.+e; 05.90.+m; 89.70.+c

Keywords: Neural networks; Phase oscillators; Random matrices

1. Introduction

The celebrated Hopfield model of associative memory [1] has provided fundamental insights into the origin of neural computations and has since stimulated much interest. In this model, neurons in the network assume discrete values (e.g., $+1$ and -1) and a set of patterns is stored such that when a new pattern is presented, the network responds by producing a stored pattern that most closely resembles the new pattern. Of interest is then the capacity, the maximum number of patterns per neuron that the network can "memorize". The physical significance

* Corresponding author. Tel.: +1 480 965 6668; fax: +1 480 965 0461.
1 Present Address: Department of Mathematics, Southern Methodist University, 208 Clements Hall, Dallas, TX 75275-0156, USA
E-mail address: yclai@chaos1.la.asu.edu (Y.-C. Lai); tnishi@smu.edu (T. Nishikawa).

0167-2780/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2004.06.011

of Hopfield's work lies in his proposal of the energy function and his idea that memories are *dynamically stable* attractors, naturally bringing concepts and tools from statistical and nonlinear physics into neuro- and information sciences as well as engineering.

Recent empirical findings in neuroscience [2,3] suggesting that synchronous firing of specific neurons is ubiquitous in the brain have stimulated several theoretical studies of models of associative memory based on the idea of temporal coding of information [4–9]. Such models typically consist of coupled oscillators interacting with each other according to a Hebbian rule, and the patterns are stored as phase-locked oscillations. One advantage of this type of model is that it can be naturally implemented using variety of oscillatory devices including phase-locked loop circuits [10], laser oscillators [11], and MEMS resonators [12].

The equation of motion for a network of coupled oscillators can be reduced to a phase model under fairly moderate conditions. Assuming that interactions are weak and that the oscillators have stable limit cycles with nearly identical periods, Kuramoto [13] has shown that the equations of motion for a network of n oscillators can be reduced to equations for the phase variables $\theta = (\theta_1, \dots, \theta_n)^{\mathrm{T}}$:

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^{n} \Gamma_{ij}(\theta_j - \theta_i), \quad i = 1, \dots, n, \tag{1}$$

where $\Gamma_{ij}(\phi)$ is a 2π-periodic function determining the coupling between neuron i and j. Reduction to a similar phase model can be obtained more rigorously (Theorems 9.1 and 9.2 in [14]) for weakly coupled oscillators. The phase model provides a convenient framework to study the phase-locking phenomenon.

In view of this reduction, many previous investigations have focused on the reduced Eq. (1) to compute the memory capacity of the model. The capacity is usually defined by the transition point at which the states that encode memory patterns *with* small amount of error become unstable (or cease to exist). It can be calculated through an extension of a standard, mean-field treatment analogous to the one used for the classical Hopfield model. It was found that an extensive number of binary patterns could be stored in a large network leading to the capacity of up to 0.042 patterns per neuron (compared to 0.138 for the Hopfield model). Even under the assumption that the network is to memorize uniformly distributed phase patterns (which may be more natural assumption than that of binary patterns), a mean-field theory can be applied to show that the capacity is about 0.038 when small amount of error is allowed in the recall process [4,15,16].

However, in many engineering applications, retrieval of the memory patterns with no error is often desired. For the classical Hopfield model, the storage of an extensive number of patterns is no longer possible, and the capacity for the error-free retrieval is known to be $1/(2\log n)$ patterns per neuron [17] for large n, where n is the total number of neurons. We have recently reported [22] results on the capacity in this sense for the oscillatory model of associative memory and the main aim of this article is to present in full detail the results obtained through rigorous mathematics using the framework of probability theory in the same spirit as in [17]. We also present additional numerical results on the basin of attraction for the retrieval solutions.

It turns out that the perfect retrieval solutions of the oscillatory models in the previous work appear to be typically *unstable* regardless of the network size, as long as the number of memory patterns is more than two [6]. This means that the capacity is $2/n$, and clearly puts oscillatory models at a disadvantage. This observation is in fact consistent with the analysis in the following sections. To overcome this malady, we will introduce a new, second-order mode in the coupling function $\Gamma_{ij}(\phi)$ and prove that this indeed results in stabilizing the solutions representing patterns. The capacity is then computed for the new model, which turns out to be at least $2\varepsilon^2/\log n$ patterns per neuron, where ε is the parameter representing the strength of the new term. Note that the capacity scales with n in a similar manner as the Hopfield case, but it can be increased by increasing the new parameter ε. However, increasing ε also tends to stabilize solutions encoding other patterns (than the memory), meaning that if ε is too large, the solution for every possible pattern becomes stable, and the system cannot distinguish the memory patterns from others. We have shown below that $\varepsilon < 1$ guarantees that the solution for a randomly chosen pattern is unstable with probability 1. We also show rigorously that $\varepsilon < 1/8$

guarantees that all symmetric mixture solutions are unstable (although they seem to remain unstable for somewhat larger ε than $1/8$ for a typical choice of p and n). In summary, the system will function (in the large system size limit) as an associative memory as long as $0 < \varepsilon < 1/8$ and at most $2\varepsilon^2/\log n$ patterns per neuron are stored.

The rest of the article is organized as follows. In Section 2, we introduce our model. Section 3 describes some properties of the energy function for the system that is useful for understanding its dynamics. Then, in Section 4, we analyze the stability of various solutions and compute the capacity of the network. We give some comments on the basin of attraction of the solution for the memory pattern in Section 5, and conclude with some remarks in Section 6.

2. The model

Many of the previous studies of weakly coupled oscillators with nearly identical frequencies have focused on the sinusoidal coupling functions to make the problem tractable. Here, we consider the coupling function $\Gamma_{ij}(\phi)$ with a second-order Fourier mode, namely,

$$\Gamma_{ij}(\phi) = C_{ij} \sin \phi + \frac{\varepsilon}{n} \sin 2\phi, \tag{2}$$

where ε is a parameter and C_{ij} the strength of coupling from oscillator j to i, which is given by Hebb's learning rule:

$$C_{ij} = \frac{1}{n} \sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu}, \tag{3}$$

where $\xi^{\mu} = (\xi_1^{\mu}, \dots, \xi_n^{\mu})^{\mathrm{T}}$, $\xi_i^{\mu} = \pm 1$ for $\mu = 1, \dots, p$, $i = 1, \dots, n$, represents a set of p patterns to be memorized. Patterns are chosen randomly, so that ξ_i^{μ} are independent and identically distributed random variables with $P(\xi_i^{\mu} = 1) = P(\xi_i^{\mu} = -1) = 1/2$. We focus on the case for which natural frequencies of all oscillators are equal (say, $\omega_i = \omega$). After the change of variable $\theta_i \to \theta_i + \omega t$, the equation of motion (1) takes the form

$$\dot{\theta}_i = \sum_{j=1}^{n} C_{ij} \sin(\theta_j - \theta_i) + \frac{\varepsilon}{n} \sum_{j=1}^{n} \sin 2(\theta_j - \theta_i). \tag{4}$$

Note that the Eq. (1) (and hence Eq. (4)) is invariant under translation by a constant, i.e., for any solution $\theta(t)$ of (1) and a constant $c \in \mathbb{R}$, $\theta(t) + (c, \dots, c)^{\mathrm{T}}$ is also a solution. This implies that there is at least one direction in the phase space in which any solution is neutrally stable. Hence, in what follows, we will not distinguish two solutions related by constant translation.

There are 2^n fixed-point solutions to Eq. (4) corresponding to all possible binary patterns of length n. Let $\eta = (\eta_1, \dots, \eta_n)^{\mathrm{T}}$ be an n-dimensional vector of 1's and -1's representing one of those binary patterns. There is a unique (up to constant translation) fixed-point solution corresponding to the pattern η, which is characterized by

$$|\theta_i - \theta_j| = \begin{cases} 0, & \text{if } \eta_i = \eta_j, \\ \pi, & \text{if } \eta_i \neq \eta_j \end{cases} \tag{5}$$

We denote this solution by $\theta(\eta)$. In the original coordinates, they are phase-locked oscillatory solutions, in which binary patterns are encoded in the locked phase deviations of the oscillators. An example of such a solution is visualized in Fig. 1 as an illustration.

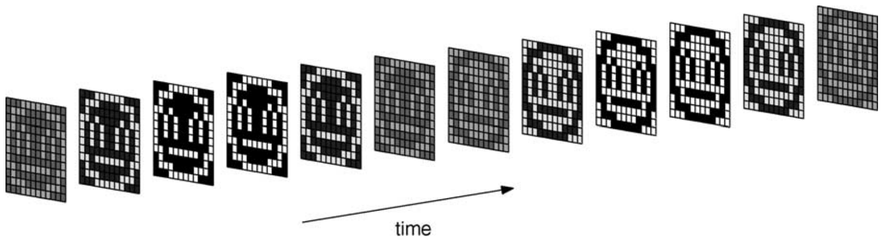

Fig. 1. A pattern encoded in the phase deviation among the oscillators. Each cell representing an oscillator is painted in gray scale according to its phase.

3. Energy function

The symmetry of the connection matrix C ensures that the system (4) can be written as a gradient system with the Lyapunov (energy) function:

$$L(\theta;\varepsilon, C) = -\frac{1}{2}\sum_{i,j=1}^{n} C_{ij}\cos(\theta_i-\theta_j) - \frac{\varepsilon}{4n}\sum_{i,j=1}^{n}\cos 2(\theta_i-\theta_j). \tag{6}$$

Thus, any solution will eventually converge to a fix point of the system located at a local minimum of the energy function (6). Using Eq. (3) for C, the energy per oscillator can be rewritten as:

$$\bar{L}(\theta;\varepsilon, \Xi) \equiv \frac{1}{n}L(\theta;\varepsilon, \Xi) = -\frac{1}{2}\sum_{\mu=1}^{p} m_\mu^2 - \frac{\varepsilon}{2}q^2, \tag{7}$$

where the $n \times p$ matrix $\Xi = (\xi^1, \ldots, \xi^p)$ and we define the order parameters:

$$m_\mu = m_\mu(\theta) = \left|\frac{1}{n}\sum_{j=1}^{n}\xi_j^\mu e^{i\theta_j}\right|, \quad \mu = 1, \ldots, p, \tag{8}$$

$$q = q(\theta)\left|\frac{1}{n}\sum_{j=1}^{n} e^{2i\theta_j}\right|. \tag{9}$$

The parameter m_μ is called the *overlap* and measures the closeness of the solution to the memory pattern ξ^μ. The parameter q measures the closeness of the solution to its nearest binary pattern. The necessity of the second term in Eq. (7) comes from the fact that a minimum of the first term is typically located near but off the fixed point corresponding to one of the patterns ξ^μ. The second term, on the other hand, always has local minima of the same depth at all fixed points representing a binary pattern. Thus, combining the two ensures that the energy minima are located precisely at the memorized patterns. We will show this more rigorously in the next section.

It is worth noting at this point that the overlaps m_μ for solutions $\theta(\eta)$ coincide with those for the Hopfield model, i.e., $m_\mu = |\sum_j \xi_j^\mu \eta_j/n|$. Moreover, since $q = 0$ for these solutions, the energy levels $\bar{L}(\eta, \Xi) \equiv L(\theta(\eta);\varepsilon, \Xi)/n$ are identical to the energy per spin in the Hopfield model and does not depend on ε. In particular, the approximate energy levels of all symmetric odd-mixture solutions are known [18], which we prove here more rigorously. The symmetric mixture solutions made of s patterns out of the p memory patterns in Ξ can be defined by the so-called "majority rule". We may just choose the first s patterns without loss of generality because of the symmetry with respect to permutation of patterns. Hence, we set

$$\eta_j^s = \mathrm{sgn}\left(\sum_{\mu=1}^{s} \xi_j^\mu\right), \tag{10}$$

where $\mathrm{sgn}\,(x) = x/|x|$. Note that the energy levels at these solutions are random variables, as they depend on random variables ξ_i^μ. Let $\xrightarrow{P}$ denote the convergence in probability, i.e., $X_n \xrightarrow{P} Y$ means that $P(|X_n - Y| > \delta) \to 0$ for every $\delta > 0$ as $n \to \infty$. The notation $a(n) = o(b(n))$ will be used to mean $a(n)/b(n) \to 0$ as $n \to \infty$.

Theorem 1. *Let $s = 2k+1$, $k = 1, 2, \ldots$ and*

$$\bar{m}_s = \frac{1}{2^{2k}}\binom{2k}{k}. \tag{11}$$

If $p = o(n)$, then, $\bar{L}(\eta^s; \Xi) \xrightarrow{P} \bar{L}_s \equiv -1/2s\bar{m}_s^2$ as $n \to \infty$.

Proof. For any $\delta > 0$, we have:

$$\begin{aligned} P\left(\left|\bar{L}(\eta^s;\Xi) + \frac{1}{2}s\bar{m}_s^2\right| \geq \delta\right) &= P\left(\left|\sum_{\mu=1}^{p} m_\mu^2 - s\bar{m}_s^2\right| \geq 2\delta\right) \\ &\leq P\left(\left|\sum_{\mu=1}^{s} m_\mu^2 - s\bar{m}_s^2\right| \geq \delta\right) + P\left(\sum_{\mu=s+1}^{p} m_\mu^2 \geq \delta\right) \\ &\leq \sum_{\mu=1}^{s} P\left(|m_\mu^2 - \bar{m}_s^2| \geq \frac{\delta}{s}\right) + P\left(\sum_{\mu=s+1}^{p} m_\mu^2 \geq \delta\right). \end{aligned} \tag{12}$$

Let $X_i^\mu = \xi_i^\mu \eta_i^s$. Writing

$$X_i^\mu = \xi_i^\mu \,\mathrm{sgn}\left(\xi_i^\mu + \sum_{\nu=1,\nu\neq\mu}^{s} \xi_j^\nu\right),$$

we see that $X_i^\mu = 1$, if $\sum_{\nu\neq\mu} \xi_i^\nu \geq 0$ and $X_i^\mu = -1$ otherwise. Note that $\sum_{\nu\neq\mu} \xi_i^\nu$ obeys the binomial distribution $\mathrm{Bin}(2k, 1/2)$. Using this fact, straightforward calculation shows that $EX_i^\mu = \bar{m}_s$. For $\mu = 1, \ldots, s$, we have $Em_\mu = \bar{m}_s$ and $\mathrm{Var}(m_\mu) = [1 + \bar{m}_s^2]/n$, and by the Chebyshev inequality, we see that for any $\delta > 0$,

$$\begin{aligned} P\left(|m_\mu^2 - \bar{m}_s^2| \geq \frac{\delta}{s}\right) &= P\left(|m_\mu + \bar{m}_s||m_\mu - \bar{m}_s| \geq \frac{\delta}{s}\right) \leq P\left(|m_\mu - \bar{m}_s| \geq \frac{\delta}{2s}\right) \\ &\leq \frac{4s^2 \mathrm{Var}(m_\mu)}{\delta^2} = \frac{4s^2(1+\bar{m}_s^2)}{\delta^2 n}, \end{aligned} \tag{13}$$

were we used $|m_\mu|,\ |\bar{m}_s| \leq 1$

For $\mu = s+1, \ldots, p$, we have $E(m_\mu^2) = 1/n$, and by the Markov inequality,

$$P\left(\sum_{\mu=s+1}^{p} m_\mu^2 \geq \delta\right) \leq \frac{E\left(\sum_{\mu=s+1}^{p} m_\mu^2\right)}{\delta} = \frac{\sum_{\mu=s+1}^{p} Em_\mu^2}{\delta} = \frac{p-s}{\delta n}. \tag{14}$$

From (12)–(14),

$$P\left(\left|\bar{L}_s(\Xi)+\frac{1}{2}s\bar{m}_s^2\right|\geq\delta\right)\leq\frac{4s^3(1+\bar{m}_s^2)}{\delta^2 n}+\frac{p-s}{\delta n}\to 0, \tag{15}$$

as $n\to\infty$. Thus, $\bar{L}(\eta^s;\Xi)\xrightarrow{P}-\frac{1}{2}s\bar{m}_s^2$. □

Note that $s=1$ leads to the solutions corresponding to memorized patterns ($\eta=\xi^\mu$). First few energy levels are: $\bar{L}_1=-1/2$, $\bar{L}_3=-3/8$, $\bar{L}_5=-45/128$, etc. One can see that there is an ordering of the energy levels: $\bar{L}_1<\bar{L}_3<\bar{L}_5<\ldots$.

4. Stability analysis

First we give general stability results that hold for any solution that corresponds to a binary pattern for a finite n. Then, in the subsections that follows, we use these results to study the stability of specific types of solutions in the limit $n\to\infty$.

The Jacobian matrix of Eq. (4) evaluated at $\theta=\theta(\eta)$ is $(2\varepsilon/n)E-2\varepsilon I+J$, where E is the $n\times n$ matrix of ones, I the $n\times n$ identity matrix, and J is defined componentwise by:

$$J_{ij}=C_{ij}\eta_i\eta_j-\delta_{ij}\sum_{k=1}^{n}C_{ik}\eta_i\eta_k=\frac{1}{n}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu\eta_i\eta_j-\frac{\delta_{ij}}{n}\sum_{k=1}^{n}\sum_{\mu=1}^{p}\xi_i^\mu\xi_k^\mu\eta_i\eta_k. \tag{16}$$

The stability of the solution $\theta(\eta)$ is determined by the eigenvalues of the Jacobian matrix. Since it is symmetric, all eigenvalues are real. Hence, the solution is stable if and only if all eigenvalues are negative.

When $\varepsilon>0$, inclusion of the second-order mode in the coupling function $\Gamma_{ij}(\phi)$ results in shifting the eigenvalues for the solution corresponding to each pattern by 2ε. This fact is used in the following theorem, which gives a condition for stability of solutions in terms of the eigenvalues of J. Let $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote the maximum and minimum eigenvalues of a matrix A, respectively.

Theorem 2. *Let η be an n-dimensional column vector of* ±1's. *The solution of* (4) *defined by* (5) *is asymptotically stable if $\lambda_{\max}(J)<2\varepsilon$, and unstable if $\lambda_{\max}(J)>2\varepsilon$.*

Proof. First note that the Jacobian $(2\varepsilon/n)E-2\varepsilon I+J$ always has eigenvalue zero associated with the eigenvector $\mathbf{1}=(1,\ \ldots,\ 1)^{\mathrm{T}}$. Let λ be a nonzero eigenvalue of J, so that $Jv=\lambda v$ for some $v\neq0$. Since J is symmetric, v is orthogonal to $\mathbf{1}$, and hence, we have $Ev=0$. Thus, we have:

$$\left[\left(\frac{2\varepsilon}{n}\right)E-2\varepsilon I+J\right]v=(\lambda-2\varepsilon)v, \tag{17}$$

which implies that $\lambda-2\varepsilon$ is an eigenvalue of the Jacobian associated with the eigenvector v. Therefore, whether $\lambda_{\max}(J)$ is above or below 2ε determines the stability of the solution defined in (5). □

Fig. 2 shows sample distributions of the $\lambda_{\max}(J)$ for several types of solutions: a memory pattern, a memory pattern with single-bit error, a symmetric three-mixture pattern, and a random pattern. Notice that even for the memory pattern $\lambda_{\max}(J)$ is always positive, indicating that the corresponding solution is unstable for $\varepsilon=0$. In fact, our numerics with various combinations of n and p suggest that the same is true[2] for any n and any $p>2$. This is indeed consistent with the observation in [6].

[2] For $p=1,2$, one can show that the solutions for the memory patterns are at least neutrally stable ($\lambda_{\max}\leq0$) by using the Gerschgorin Theorem.

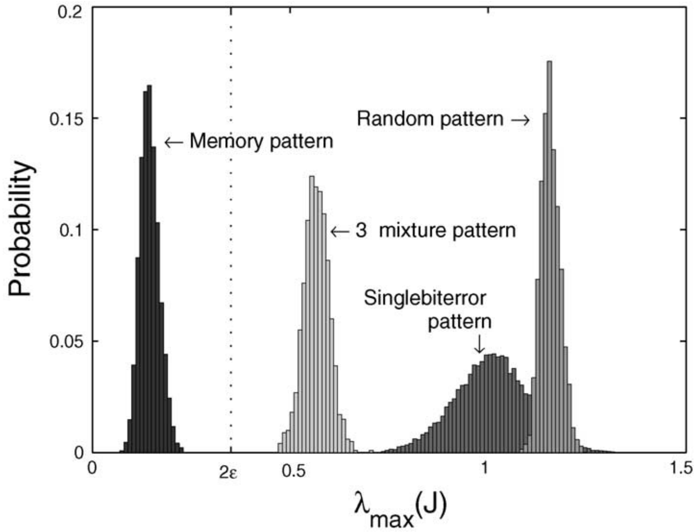

Fig. 2. The distribution of the maximum eigenvalue $\lambda_{\max}(J)$ for three types of solutions: a memory pattern, a memory pattern with single-bit error, a symmetric three-mixture pattern, and a random pattern. We chose each bit of each memory pattern to be ± 1 at random with equal probabilities. The parameters of the system were $n = 1000$ and $p = 10$. According to Theorem 2, $\lambda_{\max}(J) = 2\varepsilon$ is the borderline of stability.

Having reduced the problem of stability to the eigenvalue problem for J, we are left with finding the distribution of the eigenvalues of J. Before doing that for specific types of solutions in the subsequent subsections, let us first derive some useful properties of the eigenvalues of $\tilde{C}$, the $p \times p$ correlation matrix of the memory patterns, defined by:

$$\tilde{C}_{\mu\nu} = \frac{1}{n}\sum_{i=1}^{n} \xi_i^{\mu}\xi_i^{\nu}, \tag{18}$$

or, in matrix notation, $\tilde{C} = \frac{1}{n}\Xi^{\mathrm{T}}\Xi$. As the random variables ξ_i^{μ} are independent and take values ± 1 with probability 1/2 each, the spectra of $\tilde{C}$ obeys a law of large numbers. The following theorem is due to Girko [19] (Theorem 10.3, p. 70).

Theorem 3. *Suppose that random entries $\xi_{ij}^{(n)}$, $i = 1, \ldots, n$, $j = 1, \ldots, p_n$, of the matrix $\Xi_n = [\xi_{ij}^{(n)}]_{i=1,\ldots,n}^{j=1,\ldots,p_n}$ are independent for each n, $\gamma = \lim_{n\to\infty} p_n/n$, $\gamma > 0$, $E\xi_{ij}^{(n)} = 0$, $\mathrm{Var}(\xi_{ij}^{(n)}) = 1$. Let $F_n(x)$ denote the (random) distribution function of the eigenvalues of $\frac{1}{n}\Xi^{\mathrm{T}}\Xi$. The necessary and sufficient condition for $\lim_{n\to\infty}\sup_x |F_n(x) - F(x)| = 0$ with probability one is the modified Lindeberg condition: for every $\tau > 0$,*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{p_n}\sum_{i=1}^{n}\int_{|x|>\tau} x^2\,\mathrm{d}P(\xi_{ij}^{(n)} < \sqrt{n}x) = 0,$$

where

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = \frac{\sqrt{4\gamma x - (\gamma + x - 1)^2}\;\chi\left(\frac{\sqrt{4\gamma x}}{|\gamma+x-1|} > 1\right)}{2\pi\gamma x} + \left(1 - \frac{1}{\gamma}\right)\delta(x)\chi(\gamma \geq 1).$$

Corollary 4. *If $p = o(n)$, then all eigenvalues of $\tilde{C}$ converges to one in probability as $n \to \infty$. In particular, $\lambda_{\max}(\tilde{C}) \xrightarrow{P} 1$ and $\lambda_{\min}(\tilde{C}) \xrightarrow{P} 1$ as $n \to \infty$.*

Proof. It suffices to show that, for all $\delta > 0$, $P(\lambda_{\min}(\tilde{C}) \geq 1 - \delta) \to 0$ and $P(\lambda_{\max}(\tilde{C}) \leq 1 + \delta) \to 0$ as $n \to \infty$. For a given $\delta > 0$, choose $\gamma > 0$, so that $\gamma + 2\sqrt{\gamma} < \delta$. Choose a sequence p'_n of integers such that $\lim_n p'_n/n = \gamma$. Then, writing $p = p_n$ to explicitly show the dependence on n, we have:

$$\lim_n \frac{p_n - p'_n}{n} = -\gamma < 0, \tag{19}$$

and hence, $p_n < p'_n$ for large enough n. For each realization of Ξ, construct Ξ' by adding extra rows with random ± 1 entries, so that Ξ' has p'_n rows. Then $\tilde{C} = (1/n)\Xi^{\mathrm{T}}\Xi$ is the upper left submatrix of $\tilde{C}' = (1/n)\Xi'^{\mathrm{T}}\Xi'$, and hence, by the interlacing property of symmetric matrices (see, for example, [20]) (p. 396),

$$\lambda_{\min}(\tilde{C}') \leq \lambda_{\min}(\tilde{C}), \quad \lambda_{\max}(\tilde{C}) \leq \lambda_{\max}(\tilde{C}'). \tag{20}$$

The modified Lindeberg condition in Theorem 3 is satisfied, since the integral vanishes for large enough n. The other assumptions are clearly satisfied, and hence, the theorem can be applied to give:

$$P(\lambda_{\min}(\tilde{C}) \leq 1 - \delta) \leq P(\lambda_{\min}(\tilde{C}') \leq 1 + \gamma - 2\sqrt{\gamma}) \tag{21}$$

$$= 1 - F(1 + \gamma - 2\sqrt{\gamma}) \to 0, \tag{22}$$

$$P(\lambda_{\max}(\tilde{C}) \geq 1 + \delta) \leq P(\lambda_{\max}(\tilde{C}') \geq 1 + \gamma + 2\sqrt{\gamma}) \tag{23}$$

$$= F(1 + \gamma + 2\sqrt{\gamma}) \to 0 \tag{24}$$

as $n \to \infty$. □

A useful lower bound on $\lambda_{\max}(J)$ can be obtained by simple application of the variational principle. Let $x_i = \eta_i \xi_i^\nu$ for some fixed ν. Using (16), we have:

$$\lambda_{\max}(J) \geq \frac{\sum_{i=1}^n \sum_{j=1}^n J_{ij} x_i x_j}{\sum_{i=1}^n x_i^2} \tag{25}$$

$$= \frac{1}{n^2} \sum_{\mu=1}^p \sum_{i=1}^n \sum_{j=1}^n \xi_i^\mu \xi_j^\mu \xi_i^\nu \xi_j^\nu - \frac{1}{n^2} \sum_{\mu=1}^p \sum_{i=1}^n \sum_{k=1}^n \xi_i^\mu \xi_k^\mu \eta_i \eta_k$$

$$= \sum_{\mu=1}^p \tilde{C}_{\mu\nu}^2 - \sum_{\mu=1}^p m_\mu^2 = \sum_{\mu=1}^p \tilde{C}_{\mu\nu}^2 + 2\bar{L}(\eta, \Xi). \tag{26}$$

4.1. Perfect retrieval

The perfect retrieval solutions are defined by (5) with $\eta = \xi^\nu$, $\nu = 1, \ldots, p$. Letting $\eta = \xi^\nu$ for a fixed ν in (16), we get

$$J = S + D - \left(1 + \frac{p-1}{n}\right) I, \tag{27}$$

where

$$S_{ij} = \frac{1}{n}\sum_{\mu=1}^{p} \xi_i^\mu \xi_i^\nu \xi_j^\mu \xi_j^\nu, \tag{28}$$

$$D_{ij} = -\frac{\delta_{ij}}{n}\sum_{\mu\neq\nu}\sum_{k\neq i} \xi_i^\mu \xi_i^\nu \xi_k^\mu \xi_k^\nu. \tag{29}$$

The stability of these solutions are then determined by $\lambda_{\max}(J)$, whose statistical behavior depends on the rate of growth of $p = p_n$ with n. For evaluation of the capacity of the network, we ask how big p can be as a function of n.

Let us first consider the case $\varepsilon = 0$. In this case, the model reduces to the one previously studied. In particular, Aonishi [6] has reported that the solutions corresponding to memory patterns appear to be unstable for any combination of $p > 2$ and n. Here, we present an argument that supports this claim. According to Theorem 2, the condition for instability is $\lambda_{\max}(J) > 0$ in this case. A lower bound for $\lambda_{\max}(J)$ can be obtained by choosing x to be the normalized eigenvector of S associated with the eigenvalue $\lambda_{\max}(S)$ in (25):

$$\lambda_{\max}(J) \geq \frac{x^{\mathrm{T}} Jx}{x^{\mathrm{T}} x} = \lambda_{\max}(S) + \frac{x^{\mathrm{T}} Dx}{x^{\mathrm{T}} x} - 1 - \frac{p-1}{n}. \tag{30}$$

Taking the averages on both sides, we obtain:

$$E\lambda_{\max}(J) \geq E\lambda_{\max}(S) + E\left[\frac{x^{\mathrm{T}} Dx}{x^{\mathrm{T}} x}\right] - 1 - \frac{p-1}{n} \approx 2\sqrt{\frac{p}{n}} > 0. \tag{31}$$

Here, we used the approximation, which seems to be valid for $1 \ll p \ll n$, that the eigenvector x and the components of D are nearly independent, which leads to $E\left[x^{\mathrm{T}} Dx/x^{\mathrm{T}} x\right] \approx 0$. We also used $E\lambda_{\max}(S) \approx 1 + p/n + 2\sqrt{p/n}$, which follows from Theorem 3 if n and p are both large. This approximation appears to be reasonable, as is evident in Fig. 3 for $n = 1000$. The figure also shows that another approximation $E\lambda_{\max}(J) \approx p/n + 2\sqrt{p/n}$ seems to be valid. However, as we will see next, this approximation cannot be valid for $p, n \to \infty$ with p/n fixed, since $\lambda_{\max}(J)$ must diverge in this limit.

Let us now consider the case with arbitrary ε. The main theorem below shows that the borderline case is $p_n = O(n/\log n)$.

Theorem 5. *Let*

$$\overline{\alpha} = \limsup_n \frac{p_n \log n}{n}, \qquad \underline{\alpha} = \liminf_n \frac{p_n\ \log n}{n}, \tag{32}$$

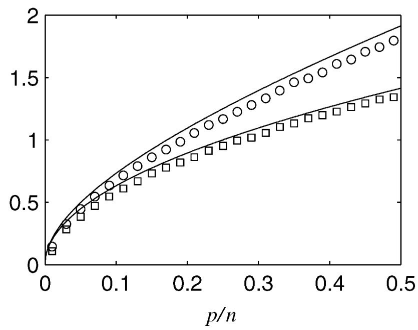

Fig. 3. The average of $\lambda_{\max}(J)$ (circles) and its lower bound $E\lambda_{\max}(S) + E[x^{\mathrm{T}} Dx/x^{\mathrm{T}} x] - 1 - (p-1)/n$ for $n = 1000$. The average is estimated with 100 realizations. The solid curves are $p/n + 2\sqrt{p/n}$ (top) and $2\sqrt{p/n}$ (bottom).

and $\varepsilon > 0$. If $\overline{\alpha} < 2\varepsilon^2$, then the solution corresponding to ξ^ν is asymptotically stable with probability one in the limit $n \to \infty$. If $\underline{\alpha} > (1+2\varepsilon)^2/2$, then it is unstable with probability tending to one as $n \to \infty$.

In other words, when n is large, $p/n < 2\varepsilon^2/\log n$ guarantees the stability, while $p/n < (1+2\varepsilon)^2/(2\log n)$ guarantees the instability of the perfect retrieval solutions. In particular, if $p_n = cn$ for some constant c, then $\underline{\alpha} = \infty$ and therefore the solution is unstable no matter how large ε is.

To prove Theorem 5, we need the following lemma.

Lemma 6. *Let $x > 0$. If $\overline{\alpha} < x^2/2$, then $P(\max_i D_{ii} \geq x) \to 0$ as $n \to \infty$. If $\underline{\alpha} > x^2/2$, then $P(\max_i D_{ii} \geq x) \to 1$ as $n \to \infty$.*

Proof. The proof uses an idea similar to that for the Big Theorem in [17]. Suppose $\overline{\alpha} < x^2/2$. Since the terms in (28) are mutually independent, a version of large deviation lemma [17] (Lemma B′) may be applied to obtain:

$$q \equiv P(D_{ii} \geq x) = P\left(-\sum_{\mu\neq\nu}\sum_{k\neq i}\xi_i^\mu\xi_i^\nu\xi_k^\mu\xi_k^\nu \geq nx\right) \sim \Phi\left(\frac{nx}{\sqrt{(p-1)(n-1)}}\right)$$
$$\sim \frac{1}{\sqrt{2\pi}}\frac{\sqrt{(p-1)(n-1)}}{nx}\exp\left(-\frac{n^2x^2}{2(p-1)(n-1)}\right) \sim \sqrt{\frac{p}{2\pi nx^2}}\exp\left(-\frac{nx^2}{2p}\right), \tag{33}$$

where $\Phi(t)$ denotes the distribution function of the standard normal random variable, and we have used the approximation formula:

$$\Phi(t) \cong \frac{1}{t\sqrt{2\pi}}e^{-t^2/2}, \tag{34}$$

valid for large t. Using $\overline{\alpha} < x^2/2$ in (33) yields:

$$q \lesssim \frac{1}{n\sqrt{4\pi\,\log n}}, \tag{35}$$

which implies that nq can be made arbitrarily small by making n large. Let $A_i = \{D_{ii} \geq x\}$ for $i = 1, 2, \ldots, n$. By Lemma C in [17], we have, for every even K such that $1 \leq K \leq n$,

$$\sum_{k=1}^{K}(-1)^{k-1}\sigma_k \leq P(\max_i D_{ii} \geq x) = P(A_1 \cup \cdots \cup A_n) \leq \sum_{k=1}^{K-1}(-1)^{k-1}\sigma_k, \tag{36}$$

where

$$\sigma_k = \sum_{j_1<\cdots<j_k} P(A_{j_1} \cap \cdots \cap A_{j_k}). \tag{37}$$

Application of Lemma 2 in [17] yields:

$$\sigma_k \sim \binom{n}{k}q^k \sim \frac{(nq)^k}{k!}. \tag{38}$$

For any given $t > 0$, take large enough n to ensure $nq < t$. Then,

$$P(\max_i D_{ii} \geq x) \lesssim \sum_{k=1}^{K-1}(-1)^{k-1}\frac{t^k}{k!} \to 1 - e^{-t} \tag{39}$$

as $K \to \infty$. Since t was arbitrary, the right-hand side can be made as small as one wishes, proving $P(\max_i D_{ii} \geq x) \to 0$ as $n \to \infty$.

Now suppose $\underline{\alpha} > x^2/2$. Proceeding similarly as above, we obtain:

$$q \sim \sqrt{\frac{p}{2\pi n x^2}} \exp\left(-\frac{nx^2}{2p}\right) \gtrsim \frac{1}{\sqrt{4\pi \log n}} \exp\left(-\frac{x^2 \log n}{2\underline{\alpha}}\right) = \frac{n^{-x^2/2\underline{\alpha}}}{\sqrt{4\pi \log n}}, \tag{40}$$

yielding

$$nq \gtrsim \frac{n^{1-x^2/2\underline{\alpha}}}{\sqrt{4\pi \log n}}, \tag{41}$$

which diverges as $n \to \infty$. Thus, for a given $t > 0$, we have $nq > t$ for large enough n, and therefore,

$$P(\max_i D_{ii} \geq x) \gtrsim \sum_{k=1}^{K} (-1)^{k-1} \frac{t^k}{k!} \to 1 - e^{-t}. \tag{42}$$

Taking $t \to \infty$ proves that $P(\max_i D_{ii} \geq x) \to 1$ as $n \to \infty$. □

Proof (Proof of Theorem 5). Suppose $\overline{\alpha} < 2\varepsilon^2$. Choose $\delta > 0$ so that $\overline{\alpha} < 2\delta^2 < 2\varepsilon^2$. Note first that the positive eigenvalues of $S = \frac{1}{n}\Xi\Xi^{\mathrm{T}}$ coincide with the eigenvalues of $\tilde{C} = \frac{1}{n}\Xi^{\mathrm{T}}\Xi$. Since

$$\lambda_{\max}(J) \leq \lambda_{\max}(S) + \lambda_{\max}(D) - 1 - \frac{p-1}{n} \leq \lambda_{\max}(\tilde{C}) + \max_i D_{ii} - 1 - \frac{p-1}{n}, \tag{43}$$

we have

$$\begin{aligned} P(\lambda_{\max}(J) \geq 2\varepsilon) &\leq P\left(\lambda_{\max}(\tilde{C}) + \max_i D_{ii} \geq 1 + \frac{p-1}{n} + 2\varepsilon\right) \\ &\leq P\left(\lambda_{\max}(\tilde{C}) \geq 1 + 2(\varepsilon - \delta) + \frac{p-1}{n}\right) + P(\max_i D_{ii} \geq 2\delta) \\ &\leq P(\lambda_{\max}(\tilde{C}) \geq 1 + 2(\varepsilon - \delta)) + P(\max_i D_{ii} \geq 2\delta). \end{aligned} \tag{44}$$

Application of Lemma 4 and Lemma 6 with $x = 2\delta$ then proves the first part of the theorem.

Now suppose $\underline{\alpha} > (1 + 2\varepsilon)^2/2$. We have

$$\lambda_{\max}(J) \geq \max_i J_{ii} = -1 + \frac{1}{n} + \max_i D_{ii}. \tag{45}$$

So,

$$P(\lambda_{\max}(J) \geq 2\varepsilon) \geq P\left(\max_i D_{ii} \geq 1 + 2\varepsilon - \frac{1}{n}\right) \geq P(\max_i D_{ii} \geq 1 + 2\varepsilon) \to 1 \tag{46}$$

as $n \to \infty$ by applying Lemma 6 with $x = 1 + 2\delta$. This completes the proof of the theorem. □

As we mentioned earlier, $\lambda_{\max}(J)$ indeed diverges with probability one if $\underline{\alpha} = \infty$, since ε can then be chosen arbitrarily large in (46).

4.2. Random patterns

Consider the case where each η_i is chosen randomly and independently to be ± 1 with equal probability, i.e., η is chosen in exactly the same fashion as ξ^μ. It is straightforward to show, using the lower bound (26) and taking the expected value of both sides, that $E\lambda_{\max}(J) \geq 1 - 1/n$. Actually, we have a stronger result.

Theorem 7. *If* $p = o(n^2)$, *then* $P(\lambda_{\max}(J) \le 1 - \delta) \to 0$ *as* $n \to \infty$ *for every* $\delta > 0$, *i.e.*, $\lambda_{\max}(J) \gtrsim 1$.

Proof. Using (26) and the Chebyshev inequality, we have:

$$\begin{aligned}
P(\lambda_{\max}(J) \le 1 - \delta) &\le P\left(\left|\sum_{\mu=1}^{p} m_\mu^2 - \sum_{\mu\neq\nu} \tilde{C}_{\mu\nu}^2\right| \ge \delta\right) \le \delta^{-2} E\left(\sum_{\mu=1}^{p} m_\mu^2 - \sum_{\mu\neq\nu} \tilde{C}_{\mu\nu}^2\right)^2 \\
&= \delta^{-2}\left[E\left(\sum_{\mu=1}^{p} m_\mu^2\right)^2 - 2E\left(\sum_{\mu=1}^{p} m_\mu^2\right)\left(\sum_{\mu\neq\nu} \tilde{C}_{\mu\nu}^2\right) + E\left(\sum_{\mu\neq\nu} \tilde{C}_{\mu\nu}^2\right)^2\right] \\
&= \delta^{-2}\left[\frac{6p + p(p-1)}{n^2} - \frac{2p(p-1)}{n^2} + \frac{6(p-1) + (p-1)(p-2)}{n^2}\right] \\
&= \frac{2(5p-2)}{n^2\delta^2} \to 0
\end{aligned} \tag{47}$$

as $n \to \infty$. □

4.3. Symmetric mixtures

We have shown in Section 3 that the energy levels of the symmetric, odd s-mixture patterns η^s converge in probability to a fixed spectrum as $n \to \infty$. Here, we show that there are lower bounds on $\lambda_{\max}(J)$ of these solutions, which converge to a fixed spectrum. By (26), we have:

$$\lambda_{\max}(J) \ge \sum_{\mu=1}^{p} \tilde{C}_{\mu\nu}^2 + 2\bar{L}_s(\Xi) \ge 1 + 2\bar{L}_s(\Xi). \tag{48}$$

Suppose that $p = o(n)$. Since $\bar{L}_s(\Xi) \xrightarrow{P} -1/2s\bar{m}_s^2$ by Theorem 1, it follows that $P(\lambda_{\max}(J) \ge 1 - s\bar{m}_s^2 - \delta) \to 1$ as $n \to \infty$, for any $\delta > 0$. In other words, $\lambda_{\max}(J)$, for large n and $p = o(n)$, is at least $\ell_s \equiv 1 - s\bar{m}_s^2$ with probability one. First few values of ℓ_s are: $\ell_1 = 0$, $\ell_3 = 1/4$, $\ell_5 = 19/64$, etc. In particular, $\lambda_{\max}(J)$ for all symmetric mixture patterns (except for $s = 1$) are above 1/4, implying that these spurious solutions will typically not be stable as long as $\varepsilon < 1/8$.

We note, however, that somewhat weaker condition on ε seems to be sufficient for ensuring the instability of the symmetric mixture solutions for a typical choice of p and n. The distribution of $\lambda_{\max}(J)$ for the symmetric three-mixture solution shown in Fig. 2 suggests that $\lambda_{\max}(J)$ is always significantly larger than 1/4, implying that ε significantly larger than 1/8 will still guarantee the instability of the spurious solutions.

5. Basin of attraction

A natural question that arises after establishing the local stability of a solution is that of the global stability. The existence of the energy function $L(\theta; \varepsilon, \Xi)$ in (6) ensures that any solution of the system converges to a phase-locked solution as $t \to \infty$. On the other hand, the local stability of the solutions representing the memory patterns means that there is an open basin of attraction for each of these solutions. How large are these basins? Put in other words, how close does the initial condition needs to be to one of the memory patterns, in order for the network to evolve into the phase-locked state that encodes that pattern?

To quantify the size of these basins, we look at the relationship between initial and final overlaps defined by (8) with initial and final θ, respectively. Recall that overlap of one corresponds to zero distance, while overlap of zero

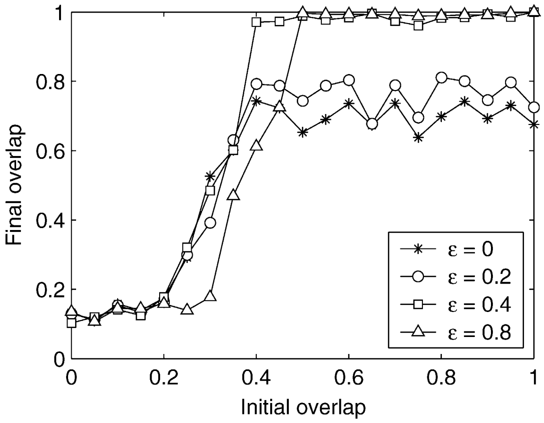

Fig. 4. The final overlap after 1000 time units as a function of the initial overlap for $n = 1000$, $p = 40$, and different values of ε (= 0, 0.2, 0.4, 0.8). The final overlap is averaged over the results from 10 random initial conditions with the same initial overlap.

indicates that patterns are as far apart as possible[3]. Fig. 4 shows typical plots of the (average) final overlap against the initial overlap.

Each plot shares a general feature: As the initial overlap decreases from one, the final overlap stays approximately constant until the initial overlap reaches the critical value, after which the final overlap drops sharply. The critical initial overlap appears to be around 0.4 for $\varepsilon = 0, 0.2, 0.4$ and around 0.5 for $\varepsilon = 0.8$. This critical value of initial overlap marks the boundary of the basin of attraction of the memory pattern solution in question, while the value of the final overlap for initial overlaps above the critical value represents the error in the retrieval process of the memory pattern. Fig. 4 illustrates two observations that can be made also for similar plots for other typical combinations of n and p.

- As ε increases from zero, the retrieval error decreases (larger final overlap) until it becomes zero (final overlap is one). Zero error seems to be achieved at a critical value $\varepsilon = \varepsilon_1$, above which the stability condition in Theorem 2 for the averaged λ_{max} is satisfied.
- As ε increases from zero, the critical value of initial overlap (the size of basin) does not change much until another critical value $\varepsilon = \varepsilon_2$ above which it increases (the basin shrinks). The transition at $\varepsilon = \varepsilon_2$ appears to correspond to the point at which the first solution other than that of a memory pattern becomes stable. This solution is usually the three-mixture solution.

Thus, in the range $\varepsilon_1 < \varepsilon < \varepsilon_2$, the optimal performance as associative memory is achieved: No error in the retrieval process and no other solution than that of memory patterns is stable, leading to maximal size of the basin of attraction for the memory pattern solutions.

In order to see the trasitions more clearly, we present another set of numerics to quantify the size of the basin in a different way. In this computation, the basin size is estimated via the probability that a randomly chosen initial condition (uniformly over the entire phase space) leads to one of the solutions corresponding to the memorized patterns. Fig. 5 shows a typical plot of such probabilities as a function of ε. It clearly shows the first transition around $\varepsilon_1 \approx 0.2$, after which almost all solutions converge to the memory pattern solutions as long as $\varepsilon < \varepsilon_2 \approx 0.6$. For

[3] A pattern and its inversion (e.g. $\{1, -1, -1\}$ and $\{-1, 1, 1\}$) cannot be distinguished in our network, as well as in the Hopfield network. Thus, such a pair is considered to be at distance zero (yielding overlap of one), while the distance between a pair is maximal when exactly half of the bits differ, leading to overlap of zero.

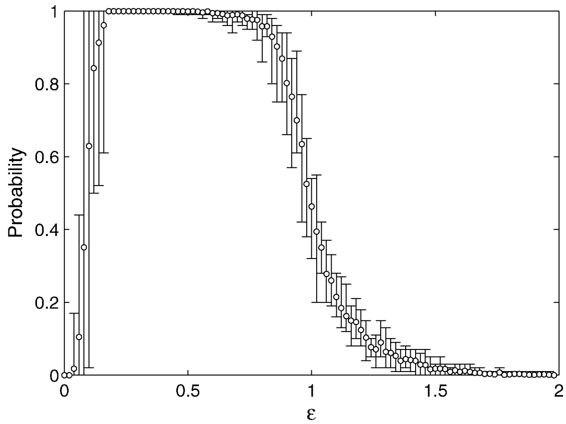

Fig. 5. The probability of perfect memory recall as a function of ε for $n = 400$, $p = 4$. The average probability over 10 realizations of memorized patterns is shown with open circles. For each realization, 100 random initial conditions are chosen to be integrated for 200 time units for the estimates of the probability that the overlap settles within 10^{-5} of one. The errorbars are drawn with the minimum and maximum estimates found in the 10 realizations.

$\varepsilon > \varepsilon_2$, the probability gradually descreases to almost zero at $\varepsilon = 2$, due to the increasing size of the basins of other solutions that become stable. By comparing Fig. 5 with Fig. 2, we see that $\varepsilon_1 \approx 0.2$ corresponds approximately to the location of the distribution of $\lambda_{\max}(J)$ for the memory patterns, and that $\varepsilon_2 \approx 0.6$ corresponds to where $\lambda_{\max}(J)$ is distributed for the three-mixture patterns. This confirms the observations we have made above that the two transition points seem to arise from the stability conditions for the memory pattern solutions and the three-mixture solutions.

6. Conclusions

In this article, we have presented a thorough analysis of the local stability of the perfect memory pattern solutions for a new type of oscillatory model of associative memory. Our model includes an extra, second-order Fourier mode in the coupling function, which enable us to control the stability of the solutions for all patterns and to distinguish the memory pattern from others by their stability. The functions $\theta_j(t)$ in our model are closely related to the cumulative distribution function of spikes in neural networks [21]. The capacity of our model turns out to follow the same scaling with the number of neurons as in the case of the classical Hopfield model, but with a prefactor that depends on the new parameter ε that control the relative strength of the two terms in the coupling function. Our conclusion is that, with a simple modification, oscillatory models of associative memory based on phase locking with a Hebbian connection scheme are capable of performing almost as well as the Hopfield model.

Our model can be modified to allow storage of patterns with n_s symbols instead of two. Similar stability results should follow in a straightforward manner simply by replacing the second term of the coupling function with $(\varepsilon/n)\sin(n_s\phi)$. More natural assumption for the coding of temporal information in terms of the phase would be to encode patterns of continuous values as $\xi^\mu = \exp(i\theta^\mu)$, where θ^μ is uniformly distributed on $[0, 2\pi)$. Unfortunately, our method of modification do not extend to such a case.

Finally, we note that the inclusion of the second term in the coupling function does not change the locality of the interactions between neurons. In fact, our coupling function can in principle be implemented using a known electric circuitry, and thus it would be feasible to implement the entire network as a network of phase-locked loops.

Acknowledgments

This work was supported by DARPA/ONR Grant N00014-01-1-0943, by NSF Grant DMS-0109001, and by AFOSR Grants F49620-01-1-0317 and F49620-03-1-0290.

References

[1] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Nat. Aca. Sci. U.S.A. 79 (8) (1982) 2554–2558.
[2] C.M. Gray, P. König, A.K. Engel, W. Singer, Oscillatory responses in cat visual-cortex exhibit inter-columnar synchronization which reflects global stimulus properties, Nature 338 (1989) 334–337.
[3] E. Vaadia, I. Haalman, M. Abeles, H. Bergman, Y. Prut, H. Slovin, A. Aertsen, Dynamics of neuronal interactions in monkey cortex in relation to behavioral events, Nature 373 (1995) 515–518.
[4] J. Cook, The mean-field theory of a Q-state neural network model, J. Phys. A: Math. Gen. 22 (1989) 2057–2067.
[5] T. Aoyagi, Network of neural oscillators for retrieving phase information, Phys. Rev. Lett. 74 (1995) 4075–4078.
[6] T. Aonishi, Phase transitions of an oscillator neural network with a standard Hebb learning rule, Phys. Rev. E 58 (4) (1998) 4865–4871.
[7] M. Yamana, M. Shiino, M. Yoshioka, Oscillator neural network model with distributed native frequencies, J. Phys. A: Math. Gen. 32 (1999) 3525–3533.
[8] T. Aonishi, K. Kurata, M. Okada, Statistical mechanics of an oscillator associative memory with scattered natural frequencies, Phys. Rev. Lett. 82 (1999) 2800–2803.
[9] M. Yoshioka, M. Shiino, Associative memory storing an extensive number of patterns based on a network of oscillators with distributed natural frequencies in the presence of external white noise, Phys. Rev. E 61 (5) (2000) 4732–4744.
[10] F.C. Hoppensteadt, E.M. Izhikevich, Pattern recognition via synchronization in phase-locked loop neural networks, IEEE Trans. Neural Networks 11 (3) (2000) 734–738.
[11] F.C. Hoppensteadt, E.M. Izhikevich, Synchronization of laser oscillators, associative memory, and optical neurocomputing, Phys. Rev. E 62 (3) (2000) 4010–4013.
[12] F.C. Hoppensteadt, E.M. Izhikevich, Synchronization of MEMS resonators and mechanical neurocomputing, IEEE Trans. Circuits Syst. I Fund. Theory Appl. 48 (2) (2001) 133–138.
[13] Y. Kuramoto, Chemical oscillations, Waves and Turbulence, Springer–Verlag, Berlin, 1984.
[14] F.C. Hoppensteadt, E.M. Izhikevich, Weakly Connected Neural Networks, Springer–Verlag, New York, 1997.
[15] T. Aoyagi, K. Kitano, Retrieval dynamics in oscillator neural networks, Neural Computation, 10 (1998) 1527–1546.
[16] S. Uchiyama, H. Fujisaka, Stability of oscillatory retrieval solutions in the oscillator neural network without lyapunov functions, Phys. Rev. E 65 (2002) 061912.
[17] R.J. McEliece, E.C. Posner, E.R. Rodemich, S.S. Venkatesh, The capacity of the Hopfield associative memory, IEEE Trans. Inform. Theory 33 (4) (1987) 461–481.
[18] D.J. Amit, Modeling Brain Function, Cambridge University Press, Cambridge, 1989.
[19] V.L. Girko, An Introduction to Statistical Analysis of Random Arrays, Int. J. Bijurcation and Chaos, (2004), in press
[20] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., The Johns Hopkins University Press, Baltimore, MD, 1996.
[21] F.C. Hoppensteadt, Modeling the cumulative distribution function of spikes in neural networks, Int. J. Bijurcation and Chaos (2004), in press.
[22] T. Nishikawa, Y.-C. Lai, F.C. Hoppensteadt, Capacity of oscillatory associative-memory networks with error-free retrieval, Phys. Rev. Lett. 92 (2004) 108101.