Effect of similarity between patterns in associative memory

Sheng-Jun Wang* and Zhou Yang
School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China
(Received 23 July 2016; revised manuscript received 31 October 2016; published 12 January 2017)

We study the stability of patterns in Hopfield networks in which a part of memorized patterns are similar. The similarity between patterns impacts the stability of these patterns, but the stability of other independent patterns is only changed slightly. We show that the stability of patterns is affected in different ways by similarity. For networks storing a number of patterns, the similarity between patterns enhances the pattern stability. However, the stability of patterns can be weakened by the similarity when networks store fewer patterns, and the relation between the stability of patterns and similarity is nonmonotonic. We present a theoretical explanation of the effect of similarity on stability using signal-to-noise-ratio analysis.

DOI: 10.1103/PhysRevE.95.012309

I. INTRODUCTION

The Hopfield attractor network is a well-known model of associative memory [1]. Based on the Hebbian rule, a number of patterns can become attractors of the network. Given an input pattern, the trained network converges to the nearest attractor, i.e., the memorized pattern which most closely resembles the input. Therefore, a Hopfield network can function as a form of associative memory [1–3]. The Hopfield model is paradigmatic in computational neuroscience research [3–7]. The concept of attractor and flow in phase space in the model are used in many studies on the dynamical principle of memory.

The model has also been studied extensively by physicists [8–19], because this model resembles the Ising model of spin glasses. The Hopfield model was modified to include more realistic properties. The effect of complex architectures on the Hopfield model has been studied [8,20–30]. The model is also extended to consider a dynamical principle of memory. For example, the training method of the Hopfield network was generalized by assigning weight to patterns [9]. The weight of patterns can be interpreted as the frequency of pattern occurrence at the input of the network. The network can learn without the catastrophic destruction of the memory [9]. The mechanism of transition between attractors was studied and compared to the transitions between the representations of space in rat's brain [10,19]. Parallel pattern processing on scale-free networks was studied [11].

In a traditional Hopfield model and various modified models, memorized patterns are selected randomly and are independent. The effect of similar patterns was studied rarely [31]. However, it is ineluctable that realistic neural systems need to memorize patterns which are similar. In experimental studies of memory, it has been shown that similarity affects the memory. Especially, similarity can affect memories in different ways [32]. The short-term memory for word sequences is impaired when acoustically similar words are used [33]. Similarity is detrimental for the ability to retain memory for faces [32,34]. Conversely, similarity among items improves visual working memory for colors [32,35]. In the Ising-like model of memory, the only mechanism of memory is the attractor dynamics. Can the attractor dynamics support that similarity affects the memory in different way?

In the present work, we modify the Hopfield model to store similar patterns along with independent patterns. The effect of similarity between memorized patterns on pattern stability is studied. Using computer simulation and analytic treatment, it will be shown that the effect differs in networks storing a lot of or fewer patterns, and the mechanism will be presented.

II. MODEL

In the Hopfield model, node states $s_i = \pm 1$, $i \in \{1,2,\ldots,N\}$ are binary. The dynamical equation of the system is [1,36]

$$s_i(t+1) = \mathrm{sgn}\left[\sum_{j=1}^{N} J_{ij} s_j(t)\right], \tag{1}$$

where J_{ij} is the entry of coupling matrix. Using the Hebbian rule,

$$J_{ij} = a_{ij} \sum_{\alpha=1}^{n} s_i^\alpha s_j^\alpha, \tag{2}$$

network memorizes n patterns which are denoted by $\{s_i^\alpha\}$ ($\alpha = 1,2,\ldots,n$). These patterns are attractors in the phase space of the network [1]. $\{a_{ij}\}$ is the adjacent matrix representing the structure of the network. $a_{ij} = 1$ if node i and node j are connected, otherwise $a_{ij} = 0$. The number of connections of a node is called the connection degree. The connection degree of the ith node is denoted by k_i, i.e., $k_i = \Sigma_{j=1}^{N} a_{ij}$. Here nodes are coupled by the Erdös-Rényi (ER) random network [37]. ER random networks are generated as follows. Starting with N disconnected nodes, we connect each of possible edges with a probability $0 < p < 1$. The degree distribution is a Poisson distribution,

$$P(k) = e^{-\langle k\rangle} \frac{\langle k\rangle^k}{k!}, \tag{3}$$

where $\langle k\rangle$ denotes the average degree of nodes in a network.

In Hopfield model patterns $\{s_i^\alpha\}$ are strings consisting of ± 1. These patterns are usually independent in order to represent different things. Here we modify the model by setting two patterns similar. The first memorized pattern $\{s_i^1\}$ is a

*wangshjun@snnu.edu.cn

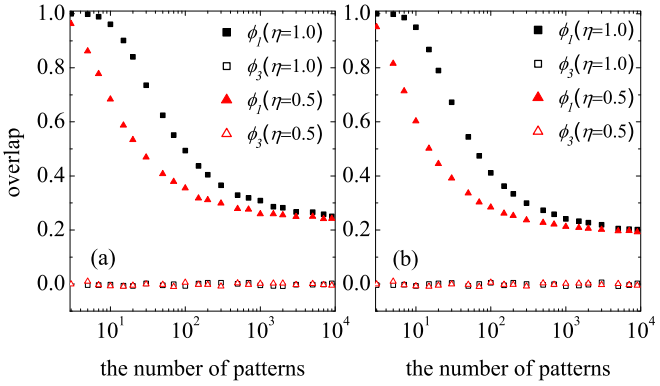

FIG. 1. The overlap between stable states and patterns versus the number of memorized patterns. (a) The first memorized pattern is used as the initial state. (b) In the initial state of networks, 10% of nodes are flipped from the first pattern. These results are averaged over 500 realizations.

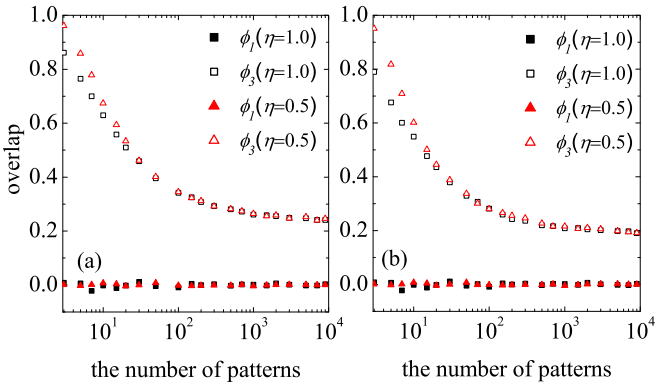

FIG. 2. The overlap between stable states and patterns versus the number of memorized patterns. (a) One of independent pattern is used as the network's initial state. (b) A spotted pattern which has 10% of nodes flipped from an independent pattern is used as the network's initial state. These results are averaged over 500 realizations.

random string, while the second pattern $\{s_i^2\}$ depends on the first one. We set $s_i^2 = s_i^1$ with probability η and $s_i^2 = -s_i^1$ with probability $1-\eta$. When $\eta = 1.0$, these two patterns are exactly the same. When $\eta = 0.5$, these two patterns are independent of each other, the same as the traditional Hopfield model.

Actually, due to the crosstalk among patterns, some errors may occur in a network even if the initial state of networks is a memorized pattern. To quantify the computational performance of Hopfield attractor networks, a memorized pattern is selected as the initial state of a network and the overlap ϕ is computed when network state has converged into a stable state [22–24]. The overlap is defined like this:

$$\phi_\alpha \equiv \frac{1}{N}\sum_{i=1}^{N} x_i s_i^\alpha, \tag{4}$$

where $x_i = \pm 1$ denotes the stable state of the ith node. When stable state is the same as the pattern, the value of overlap is 1. When they are independent, the value of overlap is 0. Additionally, when the initial state is a spotted pattern, that is, it has a portion of nodes that differ from a memorized pattern, the overlap between stable state and the pattern can be used to measure the recognition ability of Hopfield networks.

III. SIMULATION RESULTS

We first study the effect of repeated patterns on the stability. Here the second memorized pattern is the same as the first memorized pattern ($\eta = 1.0$). We take the first memorized pattern $\{s_i^1\}$ as the initial state of a network. Overlap between stable state and the first pattern is denoted by ϕ_1. We also compute overlap between stable state and independent patterns. In simulations the third memorized pattern is used to represent independent patterns, and the overlap is denoted by ϕ_3.

Figure 1(a) shows the value of overlap versus the number of patterns for both $\eta = 1.0$ and $\eta = 0.5$. From these figures one can see that the stability of the first pattern is better in networks with $\eta = 1.0$. In a large range of the number of patterns, the similarity between patterns enhances pattern stability. As the number of memorized patterns increases, the stability of patterns decreases. The overlap ϕ_1 between stable state and the first pattern is larger than 0, while the overlap ϕ_3 between stable state and other independent patterns is close to zero. The difference between ϕ_1 and ϕ_3 makes that the recognition can be performed [25].

We also consider the processing of spotted patterns in networks storing similar patterns. In our simulations, in the initial state 10% of nodes differ from the first pattern. Overlaps between stable state and the first pattern for both $\eta = 1.0$ and $\eta = 0.5$ are shown in Fig. 1(b). One can see that the value of overlap is increased by the similarity between patterns. Therefore the similarity between patterns enhances both pattern stability and recognition ability.

Next we study how similar patterns affect the stability of other independent patterns. We use one independent pattern as the initial state of Hopfield networks. For both $\eta = 1.0$ and $\eta = 0.5$, overlaps between the stable state and the third pattern are shown in Fig. 2(a). The stability of independent patterns is slightly decreased in networks storing fewer patterns and is not affected in networks storing more patterns. For recognition of spotted independent patterns, the effect of similarity between patterns is also negligibly small, as shown in Fig. 2(b). Therefore, the major effect of similarity between patterns is to change the stability of themselves.

In the following sections, we only consider pattern stability measured by setting a memorized pattern as the initial state of networks, since the stability can represent the recognition performance. The relation between overlap ϕ and similarity between patterns η is shown below.

IV. THEORETICAL RESULTS

Next we present an analytic treatment to interpret the effect of similar patterns on the stability of patterns using signal-to-noise-ratio analysis. The condition for a pattern $\{s_i\}$ to be stable can be expressed in the form of N inequalities, one for

a node i,

$$s_i \sum_{j=1}^{N} J_{ij} s_j > 0. \tag{5}$$

When we substitute the Hebbian rule, i.e., Eq. (2), into inequality (5), we get

$$s_i \sum_{\alpha=1}^{n} \sum_{j=1}^{N} a_{ij} s_i^{\alpha} s_j^{\alpha} s_j > 0. \tag{6}$$

When the stability of a memorized pattern, e.g., $s_i = s_i^{\alpha'}$, is considered, we break the left-hand side of Eq. (6) into two parts,

$$\sum_{j=1}^{N} a_{ij} s_i^{\alpha'} s_i^{\alpha} s_j^{\alpha} s_j^{\alpha'} |_{\alpha=\alpha'} + \sum_{\substack{\alpha=1 \\ \alpha \neq \alpha'}}^{n} \sum_{j=1}^{N} a_{ij} s_i^{\alpha'} s_i^{\alpha} s_j^{\alpha} s_j^{\alpha'} > 0. \tag{7}$$

The first term on the left is called the signal term [23,38], because it is equal to the connection degree k_i of node i which is positive and tends to keep node state unchanged. This term is denoted by T_s. The second part is called the noise term [23,38] because it includes $(n-1)k_i$ terms equal to $+1$ or -1 and is positive or negative randomly. This term is denoted by T_n. The stability of a pattern can be estimated by the probability of $T_s + T_n > 0$. Based on these properties of the Hopfield model, we consider that two patterns are similar in the following analysis.

A. Unstable probability of node state

1. *Similar patterns*

Here the first pattern $\{s_i^1\}$ is similar to the second pattern $\{s_i^2\}$. We take the first pattern as the initial state. The left-hand side of Eq. (6) can be broken into signal terms and noise terms like $T_{1s} + T_{2s} + T_{2n} + T_n > 0$, where subscript 1 or 2 denotes that the term is given by the first or second pattern.

Let us first derive signal terms given by the first and second pattern. For the ith node, signal term given by the first pattern is

$$T_{1s} = \sum_{j=1}^{N} a_{ij} s_i s_i^1 s_j^1 s_j. \tag{8}$$

The value of T_{1s} is equal to the degree of the ith node, $T_{1s} = k_i$. The term given by the second pattern is $\sum_{j=1}^{N} a_{ij} s_i s_i^2 s_j^2 s_j$. The probability of $s_i s_i^2 = 1$ is $P(s_i s_i^2 = 1) = \eta$, and the probability of $s_i s_i^2 = -1$ is $P(s_i s_i^2 = -1) = 1 - \eta$. Therefore, the probability of one term in the sum equal to -1 is

$$\begin{aligned} P\left(s_i s_i^2 s_j^2 s_j = -1\right) &= P\left(s_i s_i^2 = 1\right) P\left(s_j s_j^2 = -1\right) \\ &\quad + P\left(s_i s_i^2 = -1\right) P\left(s_j s_j^2 = 1\right) \\ &= 2\eta(1-\eta). \end{aligned} \tag{9}$$

In order to keep the noise term that has zero average, we take $2\eta(1-\eta)k_i$ terms of $s_i s_i^2 s_j^2 s_j = -1$ and $2\eta(1-\eta)k_i$ terms of $s_i s_i^2 s_j^2 s_j = +1$ into the noise term. So the signal term from the second pattern is

$$T_{2s} = [1 - 4\eta(1-\eta)]k_i \tag{10}$$

and the total signal term is

$$T_s = T_{1s} + T_{2s} = [2 - 4\eta(1-\eta)]k_i. \tag{11}$$

The noise term T_{2n} is the sum of $4\eta(1-\eta)k_i$ terms which are randomly equal to $+1$ or -1. So the term T_{2n} has a Gaussian distribution with standard deviation $\sigma_{2n} = \sqrt{4\eta(1-\eta)k_i}$. Besides similar patterns, all other independent patterns give noise,

$$T_n' = \sum_{\alpha=3}^{n} \sum_{j=1}^{N} a_{ij} s_i s_i^{\alpha} s_j^{\alpha} s_j. \tag{12}$$

The value of T_n' follows a Gaussian distribution with standard deviation $\sigma_n = \sqrt{(n-2)k_i}$. Therefore, the value of total noise term $T_n = T_{2n} + T_n'$ follows a Gaussian distribution of zero mean and standard deviation $\sigma = \sqrt{4\eta(1-\eta)k_i + (n-2)k_i}$, that is, the probability density is $P(T_n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(T_n)^2/2\sigma^2}$.

When $T_n < -T_s$, the ith node's state s_i evolves into $-s_i$. Therefore, the probability that nodes of degree k are unstable is

$$\begin{aligned} U_1(k) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{-[2-4\eta(1-\eta)]k} e^{-y^2/2\sigma^2} dy \\ &= \frac{1}{2}\left\{1 - \mathrm{erf}\left[\frac{(2-4\eta+4\eta^2)k}{\sqrt{2}\sigma}\right]\right\}, \end{aligned} \tag{13}$$

where $\mathrm{erf}(x) = (2/\sqrt{\pi}) \int_0^x e^{-y^2} dy$ is the error function. The index 1 denotes that the stability of the first pattern is considered.

In Fig. 3(a), we illustrate the effect of similarity on the unstable probability of similar patterns. We plot the integrand in Eq. (13), i.e., the Gaussian distribution for networks of $\eta = 0.5$ or $\eta = 1.0$. For $\eta = 0.5$, all patterns are independent. The unstable probability $U_1(k)$ is the area under the function curve in the region $(-\infty, -k)$. When $\eta = 1.0$, unstable probability $U_1(k)$ is the area in the region of $(-\infty, -2k)$. In Fig. 3(a), it is notable that the change of signal term makes

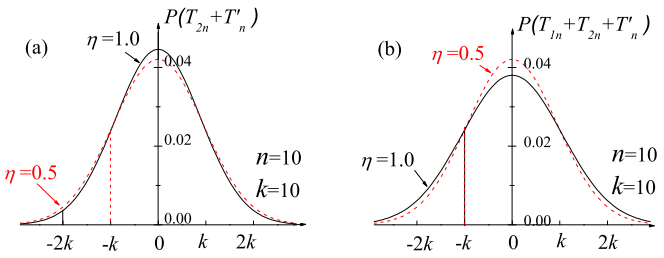

FIG. 3. Distribution of the noise term of nodes whose degree is k. (a) The stability of similar patterns is considered. (b) The stability of independent patterns is considered. The unstable probability $U_1(k)$ and $U_3(k)$ is illustrated by the area of the region under distribution curves and on the left of the vertical lines in (a) in (b). Solid lines and dashed lines represent the networks of $\eta = 1.0$ and $\eta = 0.5$, respectively. The number of patterns is $n = 10$ and the node degree is $k = 10$.

the unstable probability of similar patterns ($\eta = 1$) smaller than the unstable probability in networks storing independent patterns ($\eta = 0.5$). In other words, the stability of similar patterns becomes stronger.

2. Independent patterns

We are also interested in the effect of similarity between patterns on the stability of other independent patterns. We take one of the independent memorized patterns as network initial state. Since all these patterns are equivalent, the third pattern $\{s_i^3\}$ is used as an example of independent patterns. The signal term given by the third pattern is

$$T_s = \sum_{j=1}^{N} a_{ij} s_i s_i^3 s_j^3 s_j = k_i, \tag{14}$$

which is not changed by similarity existing in the network.

The noise term given by the first pattern is

$$T_{1n} = \sum_{j=1}^{N} a_{ij} s_i s_i^1 s_j^1 s_j, \tag{15}$$

and the noise term given by the second pattern is

$$T_{2n} = \sum_{j=1}^{N} a_{ij} s_i s_i^2 s_j^2 s_j. \tag{16}$$

These two terms are not independent. The probability that both $s_i^1 = s_i^2$ and $s_j^1 = s_j^2$ occur is η^2. The probability that both $s_i^1 = -s_i^2$ and $s_j^1 = -s_j^2$ occur is $(1-\eta)^2$. So the probability of $s_i s_i^1 s_j^1 s_j = s_i s_i^2 s_j^2 s_j$ is $\eta^2 + (1-\eta)^2$. The case that $s_i s_i^1 s_j^1 s_j = -s_i s_i^2 s_j^2 s_j$ has the probability $2\eta(1-\eta)$, which does not contribute to noise term. Therefore $T_{1n} + T_{2n}$ consists of $[\eta^2 + (1-\eta)^2]k_i$ terms equal to $+2$ or -2. This noise given by similar patterns has a Gaussian distribution with standard deviation $\sigma_{1n+2n} = \sqrt{4[\eta^2 + (1-\eta)^2]k_i}$.

Noise given by all other patterns is

$$T_n' = \sum_{\alpha=4}^{n} \sum_{j=1}^{N} a_{ij} s_i s_i^\alpha s_j^\alpha s_j. \tag{17}$$

The value of total noise term $T_n = T_{1n} + T_{2n} + T_n'$ follows a Gaussian distribution with mean value $\mu = 0$ and standard deviation $\sigma = \sqrt{4[\eta^2 + (1-\eta)^2]k_i + (n-3)k_i}$, that is, the probability density is $P(T_n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(T_n)^2/2\sigma^2}$.

When $T_n < -T_s$, the ith node's state s_i evolves into $-s_i$. Therefore, the probability that nodes of degree k are unstable is

$$\begin{aligned} U_3(k) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{-k} e^{-y^2/2\sigma^2} dy \\ &= \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{k}{\sqrt{2}\sigma}\right)\right]. \end{aligned} \tag{18}$$

The index 3 denotes that the stability of the third pattern is considered.

In Fig. 3(b), we illustrate the unstable probability of independent patterns. For different values of η, the unstable probability $U_3(k)$ is the area under the curve in the region $(-\infty, -k)$. The similarity between patterns only affects the noise term. When η is larger, standard deviation σ is larger, and unstable probability is larger, but the change is slight.

Based on aforementioned simulation and analysis, we can know that the major effect of similarity between patterns is to change the stability of similar patterns. In the following analysis, we focus on the stability of only similar patterns.

B. Overlaps between stable state and pattern

The stability of memorized patterns can be predicted by unstable probability of node state $U(k)$. The overlap ϕ between stable state and initial pattern can be written as [23]

$$\phi = \frac{1}{N}(N - 2N_{\mathrm{flip}}). \tag{19}$$

For similar patterns, the number of flipped nodes of a network is

$$\begin{aligned} N_{1\mathrm{flip}} &= \sum_k N P(k) U_1(k) \\ &= \sum_k N e^{-\langle k\rangle} \frac{\langle k\rangle^k}{k!} \frac{1}{2}\left\{1 - \mathrm{erf}\left[\frac{(2 - 4\eta + 4\eta^2)k}{\sqrt{2}\sigma}\right]\right\}. \end{aligned} \tag{20}$$

Thus overlap ϕ_1 is

$$\phi_1 = 1 - \sum_k e^{-\langle k\rangle} \frac{\langle k\rangle^k}{k!} \left\{1 - \mathrm{erf}\left[\frac{(2 - 4\eta + 4\eta^2)k}{\sqrt{2}\sigma}\right]\right\}, \tag{21}$$

where $\sigma = \sqrt{4\eta(1-\eta)k_i + (n-2)k_i}$. When parameters n and $\langle k\rangle$ are given, the value of overlap in Eq. (21) can be numerically computed for different η.

In this definition of overlap, the number of nodes which are not stable in the stored pattern is used to estimate the overlap. When these nodes flip their state, the stability of each node also changes, and overlap can be predicated using the new pattern again. The value of overlap of a network converge to a fixed value until the stable state is achieved. Therefore, the overlap defined here does not consider the cascade of the flips and is not the same as the overlap obtained in simulations. But the analysis can reflect the main mechanism of the change of the stability of patterns.

C. Overlap versus similarity

Let us investigate how the overlap between stable state and memorized pattern depends on the value of similarity η. In computer simulations, one of similar memorized patterns is used as network initial state. The average degree of networks is $\langle k\rangle = 10$. Different values of the number of patterns are used in simulations. In Fig. 4 we plot the relation between overlap and similarity η. One can see that as the value of η increases the overlap increases when the number of patterns n is large ($n = 15$ or 20). The simulation results (points) agree with the theoretical results (lines).

However, in networks storing fewer patterns, we obtain nonmonotonic relations between overlap and similarity η in simulations. It is notable that similarity between memorized patterns can weaken the stability of patterns when a network has a lower storage load. Simulation results in networks of

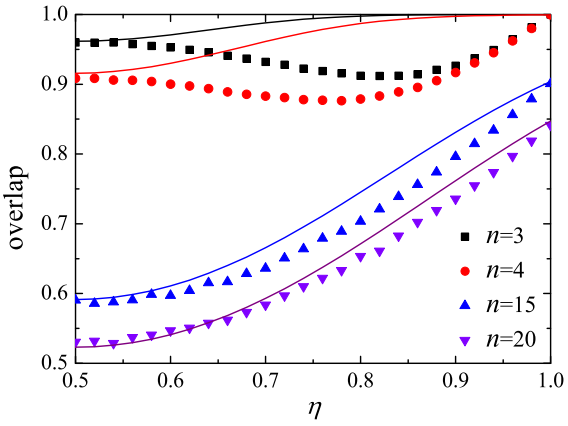

FIG. 4. Theoretical (lines) and simulation (points) results of overlap versus the similarity η for different number of patterns n. Data points are averaged over 500 realizations.

lower load do not agree with above theoretical results. We give analytic treatment of this case below.

D. Networks with low storage load

In Sec. IV A, we used the same form of signal T_{2s} from the second pattern when we considered the stability of every node, that is, signal terms are the mean value which is averaged over all node with the same degree. In networks with lower storage load, we discuss nodes with $s_i^1 = s_i^2$ (type I) and nodes with $s_i^1 = -s_i^2$ (type II) separately.

For nodes with $s_i^1 = s_i^2$ (type I), the probability of $s_i^1 s_i^2 s_j^2 s_j^1 = +1$ is

$$P\left(s_i^1 s_i^2 s_j^2 s_j^1 = +1\right) = P\left(s_j^2 s_j^1 = 1\right) = \eta, \tag{22}$$

and the probability of $s_i^1 s_i^2 s_j^2 s_j^1 = -1$ is

$$P\left(s_i^1 s_i^2 s_j^2 s_j^1 = -1\right) = P\left(s_j^2 s_j^1 = -1\right) = 1 - \eta. \tag{23}$$

We take $(1-\eta)k_i$ terms of $s_i^1 s_i^2 s_j^2 s_j^1 = +1$ and $(1-\eta)k_i$ terms of $s_i^1 s_i^2 s_j^2 s_j^1 = -1$ into noise. Then the signal from the second pattern consists of $(\eta - (1-\eta))k_i$ terms equal to $+1$. Thus signal term is

$$T_{2s}^{+} = [\eta - (1-\eta)]k_i, \tag{24}$$

where superscript "+" denotes type I nodes. This value of T_{2s}^{+} is positive. Therefore the signal term of type I nodes is increased when patterns are similar. The sum of signal is

$$T_s^{+} = T_{1s} + T_{2s}^{+} = 2\eta k_i. \tag{25}$$

For nodes with $s_i^1 = -s_i^2$ (type II), the probability of $s_i^1 s_i^2 s_j^2 s_j^1 = +1$ is

$$P\left(s_i^1 s_i^2 s_j^2 s_j^1 = +1\right) = P\left(s_j^2 s_j^1 = -1\right) = 1 - \eta, \tag{26}$$

and the probability of $s_i^1 s_i^2 s_j^2 s_j^1 = -1$ is

$$P\left(s_i^1 s_i^2 s_j^2 s_j^1 = -1\right) = P\left(s_j^2 s_j^1 = +1\right) = \eta. \tag{27}$$

We take $(1-\eta)k_i$ terms of $s_i^1 s_i^2 s_j^2 s_j^1 = +1$ and $(1-\eta)k_i$ terms of $s_i^1 s_i^2 s_j^2 s_j^1 = -1$ into noise. So the signal from the second pattern consists of $(\eta - (1-\eta))k_i$ terms equal to -1,

$$T_{2s}^{-} = -[\eta - (1-\eta)]k_i, \tag{28}$$

where superscript "−" denotes type II nodes. The value of T_{2s}^{-} is negative. Therefore the signal term of type II nodes becomes small due to the similarity. The total signal is

$$T_s^{-} = T_{1s} + T_{2s}^{-} = 2(1-\eta)k_i. \tag{29}$$

The noise term given by the second pattern is denoted by T_{2n}. The term includes $2(1-\eta)k_i$ terms equal to $+1$ or -1 for both type I and type II nodes. The total noise is

$$T_n^{\text{few}} = T_{2n} + T_n', \tag{30}$$

where $T_n' = \sum_{\alpha=3}^{n}\sum_{j=1}^{N} a_{ij} s_i s_i^\alpha s_j^\alpha s_j$. Here T_n^{few} follows the Gaussian distribution with standard deviation $\sigma^{\text{few}} = \sqrt{2(1-\eta)k_i + (n-2)k_i}$, and the probability density is $P(T_n^{\text{few}}) = \frac{1}{\sqrt{2\pi}\sigma^{\text{few}}} e^{-(T_n^{\text{few}})^2/2(\sigma^{\text{few}})^2}$.

The unstable probability of type I nodes with degree k is

$$\begin{aligned} U^{+}(k) &= \frac{1}{\sqrt{2\pi}\sigma^{\text{few}}}\int_{-\infty}^{-2\eta k} e^{-y^2/2(\sigma^{\text{few}})^2} dy \\ &= \frac{1}{2}[1 - \text{erf}(x_1)], \end{aligned} \tag{31}$$

where $x_1 = \frac{\sqrt{2}\eta k}{\sigma^{\text{few}}}$. If signal is given by only the first pattern, then an unstable probability of nodes of degree k is

$$\begin{aligned} U_{1s}(k) &= \frac{1}{\sqrt{2\pi}\sigma^{\text{few}}}\int_{-\infty}^{-k} e^{-y^2/2(\sigma^{\text{few}})^2} dy \\ &= \frac{1}{2}[1 - \text{erf}(x_2)], \end{aligned} \tag{32}$$

where $x_2 = \frac{k}{\sqrt{2}\sigma^{\text{few}}}$. So unstable probability of type I nodes given by the second pattern is

$$\begin{aligned} \Delta U^{+} &= U^{+}(k) - U_{1s}(k) \\ &= \tfrac{1}{2}[\text{erf}(x_2) - \text{erf}(x_1)]. \end{aligned} \tag{33}$$

The value of ΔU^{+} is negative and the stability of type I nodes is enhanced. We illustrate the difference using the error function in Fig. 5.

The unstable probability of type II nodes with degree k is

$$\begin{aligned} U^{-}(k) &= \frac{1}{\sqrt{2\pi}\sigma^{\text{few}}}\int_{-\infty}^{-2(1-\eta)k} e^{-y^2/2(\sigma^{\text{few}})^2} dy \\ &= \frac{1}{2}[1 - \text{erf}(x_3)], \end{aligned} \tag{34}$$

where $x_3 = \frac{\sqrt{2}(1-\eta)k}{\sigma^{\text{few}}}$. The unstable probability of type II nodes given by the second pattern is

$$\begin{aligned} \Delta U^{-} &= U^{-}(k) - U_{1s}(k) \\ &= \frac{1}{2}[\text{erf}(x_2) - \text{erf}(x_3)]. \end{aligned} \tag{35}$$

The value of ΔU^{-} is positive and the stability of type II nodes is weakened.

In Fig. 5(a), we can see that when the number of patterns is small ($n = 3$), the error function is saturated at x_2, that is, the first pattern already makes the node state stable. The value of $|\Delta U^{+}|$ is small, while ΔU^{-} is large. So type I nodes cannot

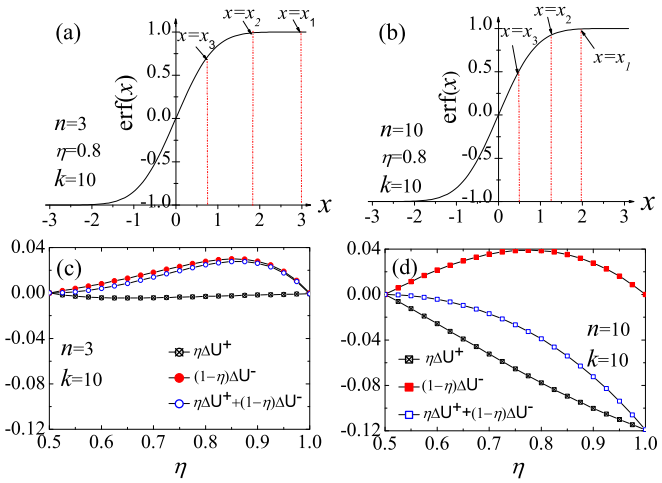

FIG. 5. Panels (a) and (b) illustrate the error functions for networks storing $n = 3$ and $n = 10$ patterns, respectively. The similarity is $\eta = 0.8$. Panels (c) and (d) are the variation of unstable probability contributed by the second pattern in networks storing $n = 3$ and $n = 10$ patterns, respectively. The node degree is $k = 10$.

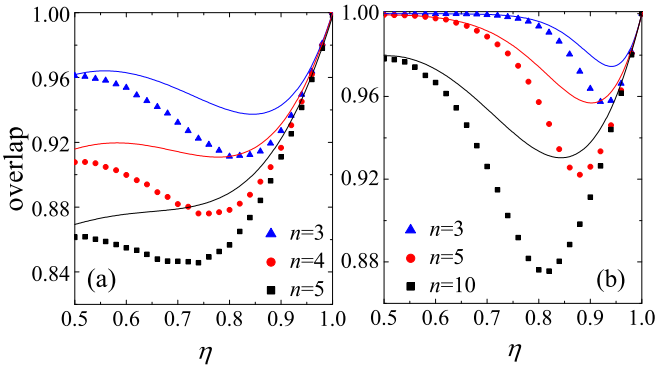

FIG. 6. The overlap versus the similarity for networks storing different number of patterns. The scatter plots are the simulation results of overlap; the curves are the theoretical results of overlap. Data points are averaged over 500 network realizations. The average degree of networks is $\langle k\rangle = 10$ and 50 in panels (a) and (b), respectively.

become more stable, and type II nodes become less stable. Multiplying by the fraction of type I and II nodes, the value of $(1-\eta)\Delta U^-$ dominates the effect of the second pattern, as shown in Fig. 5(c). Therefore the stability of the whole pattern can be decreased, and the stability can be nonmonotonic as η increases. When the number of patterns is large ($n = 10$) as shown in Fig. 5(b), error function is not saturated at x_2, and type I nodes can become more stable due to the similarity of patterns. After weighted by the fraction of type I and II nodes, the value of $\eta\Delta U^+$ is always larger, as shown in Fig. 5(d). So the pattern stability in networks storing a lot of patterns is always increased as the value of η increases.

The number of flipped nodes in type I node is

$$N^+_{\text{flip}} = \sum_k \eta N P(k) U^+(k) = \frac{\eta N}{2} \sum_k e^{-\langle k\rangle} \frac{\langle k\rangle^k}{k!} [1 - \text{erf}(x_1)]. \quad (36)$$

The number of flipped node in type II node is

$$N^-_{\text{flip}} = \sum_k (1-\eta) N P(k) U^-(k) = \frac{(1-\eta)N}{2} \sum_k e^{-\langle k\rangle} \frac{\langle k\rangle^k}{k!} [1 - \text{erf}(x_3)]. \quad (37)$$

Thus overlap ϕ_1^{few} is

$$\phi_1^{\text{few}} = \frac{1}{N}(N - 2N^+_{\text{flip}} - 2N^-_{\text{flip}}). \quad (38)$$

We use this analytic result to compute the relation between overlap and similarity η, which is shown in Fig. 6. We also show simulation results on networks with $\langle k\rangle = 10$ or $\langle k\rangle = 50$. The first pattern is used as network's initial state. In Fig. 6, theoretical results of overlap ϕ_1^{few} (lines) are also nonmonotonic with similarity η. Theoretical results qualitatively agree with simulation results (points).

Therefore the effect of similarity on type I and type II nodes differs. Similarity strengthens type I nodes and weakens type II nodes, separately. When networks have a low storage load, the stability of type I nodes is saturated, and the type II node induces that the similarity can weaken the stability of memorized patterns. Because the definition of the overlap does not include the cascade of errors in patterns, the analytic results cannot quantitatively equal to the simulation results. The analytic results provides the qualitative explanation of the effect of similarity on the stability of patterns.

V. CONCLUSIONS

The effect of similarity of memorized patterns on patterns' stability in Hopfield networks is studied. In this work two patterns are similar and other patterns are independent. We show that the stability of these two patterns is affected by similarity between them, while independent patterns are affected slightly. Similarity between patterns affects pattern stability in different ways that depend on the storage load of networks. Similarity between patterns enhances the stability of similar patterns in networks of higher load, that is, networks memorize more patterns. In the network of lower load, i.e., a few of patterns are stored in networks, pattern stability can be weakened by the similarity between patterns and stability nonmonotonically changes with similarity. The mechanism of these effects is explained by the signal-to-noise-ratio analysis.

The properties of the Hopfield model arise from the nature of the flow in phase space and do not strongly depend on precise details of modeling. The attractor dynamics can be used to understand more complex neural systems and has been used in studies on biological realistic neural networks [4]. Therefore, our results in the modeling study may provide new insight into the mechanism of associative memory in networks processing patterns which are similar.

ACKNOWLEDGMENTS

This work was supported by NSFC (Grants No. 11305098 and No. 11675096), the Natural Science Basic Research Plan in Shaanxi Province of China (Grant No. 2014JQ1028), and the Interdisciplinary Incubation Project of Shaanxi Normal University (Grant No. 5).

[1] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).
[2] M. Shanahan, Philos. Trans. R. Soc. Lond. B **367**, 2704 (2012).
[3] D. J. Amit, *Modeling Brain Functions: The World of Attractor Neural Networks* (Cambridge University Press, Cambridge, 1989).
[4] B. E. Pfeiffer and D. J. Foster, Science **349**, 180 (2015).
[5] A. Alemi, C. Baldassi, N. Brunel, and R. Zecchina, PLoS Comput. Biol. **11**, e1004439 (2015).
[6] N. Brunel, Nat. Neurosci. **19**, 749 (2016).
[7] M. I. Rabinovich, P. Varona, A. I. Selverston, and H. D. I. Abarbanel, Rev. Mod. Phys. **78**, 1213 (2006).
[8] H. Oshima and T. Odagaki, Phys. Rev. E **76**, 036114 (2007).
[9] I. Karandashev, B. Kryzhanovsky, and L. Litinskii, Phys. Rev. E **85**, 041925 (2012).
[10] R. Monasson and S. Rosay, Phys. Rev. Lett. **115**, 098101 (2015).
[11] P. Sollich, D. Tantari, A. Annibale, and A. Barra, Phys. Rev. Lett. **113**, 238106 (2014).
[12] M. C. Diamantini and C. A. Trugenberger, Phys. Rev. E **89**, 052138 (2014).
[13] A. Braunstein, A. Ramezanpour, R. Zecchina, and P. Zhang, Phys. Rev. E **83**, 056114 (2011).
[14] H. Huang, Phys. Rev. E **81**, 036104 (2010).
[15] E. Agliari, A. Barra, A. Galluzzi, F. Guerra, D. Tantari, and F. Tavani, Phys. Rev. Lett. **114**, 028103 (2015).
[16] H. Zhao, Phys. Rev. E **70**, 066137 (2004).
[17] T. Jin and H. Zhao, Phys. Rev. E **72**, 066111 (2005).
[18] E. Agliari, A. Barra, S. Bartolucci, A. Galluzzi, F. Guerra, and F. Moauro, Phys. Rev. E **87**, 042701 (2013).
[19] R. Monasson and S. Rosay, Phys. Rev. E **89**, 032803 (2014).
[20] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. **55**, 1530 (1985).
[21] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985).
[22] P. N. McGraw and M. Menzinger, Phys. Rev. E **68**, 047102 (2003).
[23] S. J. Wang, Z. G. Huang, X. J. Xu, and Y. H. Wang, Eur. Phys. J. B **86**, 424 (2013).
[24] B. M. Forrest, J. Phys. (France) **50**, 2003 (1989).
[25] D. Stauffer, A. Aharony, L. da Fontoura Costa, and J. Adler, Eur. Phys. J. B **32**, 395 (2003).
[26] I. P. Castillo, B. Wemmenhove, J. P. L. Hatchett, A. C. C. Coolen, N. S. Skantzos, and T. Nikoletopoulos, J. Phys. A **37**, 8789 (2004).
[27] J. I. Perotti, F. A. Tamarit, and S. A. Cannas, Physica A **371**, 71 (2006).
[28] Y. Bar-Yam and I. R. Epstein, Proc. Natl. Acad. Sci. USA **101**, 4341 (2004).
[29] S. J. Wang, A. C. Wu, Z. X. Wu, X. J. Xu, and Y. H. Wang, Phys. Rev. E **75**, 046113 (2007).
[30] E. K. Bunde, J. Phys. (France) **51**, 1797 (1990).
[31] F. A. Tamarit and E. M. F. Curado, J. Stat. Phys. **62**, 473 (1991).
[32] Y. V. Jiang, H. J. Lee, A. Asaad, and R. Remington, Psychon. Bull. Rev. **23**, 476 (2016).
[33] A. D. Baddeley, Q. J. Exp. Psychol. **18**, 302 (1966).
[34] M. A. Cohen, T. Konkle, J. Y. Rhee, K. Nakayama, and G. A. Alvarez, Proc. Natl. Acad. Sci. USA **111**, 8955 (2014).
[35] P. H. Lin and S. J. Luck, Vis. Cogn. **17**, 356 (2009).
[36] D. Kleinfeld, Proc. Natl. Acad. Sci. USA **83**, 9469 (1986).
[37] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Phys. Rep. **424**, 175 (2006).
[38] P. Pérez and D. Salinas, Phys. Rev. E **50**, 4182 (1994).