PERGAMON

Neural Networks 15 (2002) 69–84

Neural Networks

www.elsevier.com/locate/neunet

Contributed article

The Bifurcating Neuron Network 2: an analog associative memory

Geehyuk Lee, Nabil H. Farhat*

Electrical Engineering Department, University of Pennsylvania, 200 S 33rd Street, Philadelphia, PA 19104, USA

Received 26 September 2000; accepted 26 June 2001

Abstract

The Bifurcating Neuron (BN), a chaotic integrate-and-fire neuron, is a model of a neuron augmented by coherent modulation from its environment. The BN is mathematically equivalent to the sine-circle map, and this equivalence relationship allowed us to apply the mathematics of one-dimensional maps to the design of a BN network. The study of the bifurcating diagram of the BN revealed that the BN, under a suitable condition, can function as an amplitude-to-phase converter. Also, being an integrate-and-fire neuron, it has an inherent capability to function as a coincidence detector. These two observations led us to the design of the BN Network 2 (BNN-2), a pulse-coupled neural network that exhibits associative memory of multiple *analog patterns*. In addition to the usual dynamical properties as an associative memory, the BNN-2 was shown to exhibit *volume-holographic* memory: it switches to different pages of its memory space as the frequency of the coherent modulation changes, meaning context-sensitive memory.

Keywords: Bifurcating Neuron; Integrate-and-fire neuron; Time coding; Coherence; Analog associative memory; Bifurcating Neuron Network; Neural network

1. Introduction

The realization of associative memory in a pulse-coupled neural network (PCNN) has been a topic of interest among many scientists, and several PCNN models realizing associative memory have been proposed (Bibbig, Wennekers & Palm, 1995; Fukai & Shiino, 1995; Gerstner & van Hemmen, 1992; Herz, Li & van Hemmen, 1991; Maass & Natschlaeger, 1998; Wennekers, Sommer & Palm, 1995). Although each of the PCNN models is different and states its own unique idea, we can identify two major categories into which most of the PCNN models can be classified. The models in the first category utilize the incoherent dynamics of the PCNN (Gerstner & van Hemmen, 1992; Herz et al., 1991). In a PCNN operating in an incoherent mode, where the individual firing time of a neuron does not count, the function of a spiking neuron is similar to that of a sigmoidal neuron. In this respect, the models of the first category can be regarded as PCNN implementations of the continuous time version of the Hopfield network. The models in the second category utilize the coherent dynamics of the PCNN (Bibbig et al., 1995; Maass & Natschlaeger, 1998; Wennekers et al., 1995). In a PCNN operating in a coherent mode, the firing of a neuron is confined to small periodic windows, and the probability of the firing is determined by the firings of presynaptic neurons in the previous windows. Conceptually, the dynamical picture of a PCNN operating in such a coherent mode is akin to that of the discrete time version of the Hopfield network. The associative memory dynamics of a PCNN, whether it is coherent or incoherent, appear to be equivalent to that of the Hopfield network, of course, not in a strict mathematical sense, but in a conceptual sense. A clear consequence of such equivalence is that the PCNN models end up being a *binary* associative memory, i.e. an associative memory that can store binary patterns.

The name ‘binary associative memory’ may sound somewhat unfamiliar since it seldom appears in neural network literature. It seems that the ‘binary’ part is usually implied and understood, and there is no need to mention it explicitly. This convention seems to have a long history dating back to when Cragg and Temperley (1955) introduced an associative memory mechanism where memory resides in the hysteresis of the domain patterns of a magnetic system. The idea is extended by Caianiello (1961) by incorporating a Hebbian learning, and is taken up by Little and Shaw (1975) and again by Hopfield (1982). This means that most associative memory networks introduced so far are the offspring of the common idea based on the analogy with a magnetic system, and this explains why the modifier ‘binary’ is usually assumed. This observation naturally

* Corresponding author. Tel.: +1-215-898-5882; fax: +1-215-572-2068.
E-mail address: farhat@ee.upenn.edu (N.H. Farhat).

0893-6080/02/$ - see front matter
PII: S0893-6080(01)00100-9

Nomenclature

List of symbols

$x_i(t)$	The internal potential of BN i
$\theta_i(t)$	The threshold level of BN i
$\rho_i(t)$	The relaxation level of BN i
$\rho(t)$	Same as $\rho_i(t)$ ($\rho_i(t)$ is independent of i)
ρ_0	The amplitude of $\rho(t)$
f	The frequency of $\rho(t)$
c_i	The constant buildup rate of $x_i(t)$
$y_i(n)$	The output of BN i
$t_i(n)$	The n-th firing time of BNi
$\delta_i(t)$	The perturbation in $\theta_i(t)$ induced by presynaptic BNs
β	The reciprocal time constant of the threshold level of a BN
$u_i(n)$	The external (extrinsic) input to BN i
$u_i^f(n)$	The network (intrinsic) input to BN i
w_{ij}^k	The binary weight of the k-th connection from BN j to BN i
τ_{ij}^k	The time-delay of the k-th connection from BN j to BN i
γ	A constant parameter that controls the overall efficiency of external inputs
d	A constant parameter that controls the overall efficiency of network inputs
ξ_i^k	The i-th component of the k-th training pattern

leads to the following question: is it impossible to realize an analog associative memory in a neural network?

It seems that studies in neuroscience do not provide any evidence in favor of a binary or an analog associative memory. Although there are several biologically inspired models for a binary associative memory, as some examples are shown above, this does not disprove the possibility of analog associative memory in the brain. In fact, we have reason to believe, though we cannot prove, that the brain is utilizing analog associative memory. First, both the input (sensory signals) to and the output (motor signals) from the brain are in a graded-valued (analog) form. This means that analog associative memory can be a more useful form of memory than binary one. Second, unlike a sigmoidal neuron, which is basically a thresholding element, a real biological neuron has far more capability than simply producing all or none output. For instance, as we will mention again below, there are many neurophysiological findings that suggest information coding in the spiking time (firing instances) of a neuron. Third, it is known that a neuron in the brain typically makes synaptic connections of the order of 10^4 to its neighbors and often makes multiple connections to the same target neuron (Kuffler, Nicholls & Martin, 1992). It is hard to understand the purpose of such recurring connections[1] if we try to understand it with the binary associative memory mechanism in mind.

In the following, we will describe the Bifurcating Neuron (BN), our neuron model, and then introduce the Bifurcating Neuron Network 2 (BNN-2), a network of BNs that demonstrates a possible mechanism of analog associative memory in a PCNN. To begin with, we will summarize briefly some of the recent neurological findings that have influenced and guided our conceptualization of the BN and the BNN-2.

The so-called coding problem (Rieke, Warland, de Ruyter van Steveninck & Bialek, 1996), which deals with how the brain encodes information in neuronal spikes, has been a topic of continuing debate in neuroscience. The most widely accepted coding scheme is called rate coding, where information is represented by the mean firing rate of a neuron, whether in a temporal sense or in a spatial sense. This hypothesis was first proposed by Adrian (1926) from the study of the relationship between the activity pattern of a stretch receptor in a frog muscle and the amount of weight applied. On the other hand, recent experimental studies are revealing a growing number of new facts beyond the explanation of rate coding and are suggesting the possibility of information coding in the precise timing of neuronal spikes, namely, time coding. An especially descriptive example supporting time coding in the brain is provided by O'Keefe and Recce (1993), who studied the firing behavior of hippocampal place cells (O'Keefe & Dostrovsky, 1971). They showed that the firing phases of the place cells with respect to the theta rhythm have a high level of correlation with the animal's location on a linear runway. This result suggests the possibility of coding spatial location information in the relative firing phases of neuronal spikes.

Another topic in neuroscience that has had influence on our network design is the role of coherent activities in the brain, especially those in the gamma-band centered around 40 Hz. Some of the early observations of gamma-band oscillatory activities were made in the olfactory bulb and

[1] Recurring interconnections mean multiple connections from a presynaptic neuron to a postsynaptic neuron.

cortex of the rabbit (Freeman & Skarda, 1985), in the olfactory systems of the cat and the rat (Bressler & Freeman, 1980), in a variety of structures of the cat brain (Basar, 1983), in the cat primary visual cortex (Eckhorn, Bauer, Jordan, Brosch, Kruse, Munk et al., 1988; Gray & Singer, 1989), in the monkey visual cortex (Freeman & van Dijk, 1987), and in EEG recordings from the human skull above association and motor areas (Krieger & Dillbeck, 1987). The observation of a synchronous activity in the cat visual cortex by Gray and Singer (1989) has drawn and continues to draw special attention because, in their experiments, the synchronous activity was stimulus specific and was observed across cortical regions, e.g. across multiple visual association areas, with small phase differences. Gray and Singer related their result to the so called feature-binding hypothesis (Milner, 1974; von der Malsburg, 1981), which states that synchrony provides a means to bind together in time the features that represent a particular stimulus.

Although the role of coherence is still the focus of debate, it was necessary to define a role of coherence that can guide our network design. The one that we chose over other possibilities states that a coherent activity in an integrate-and-fire neuron network can provide a time reference. This role may sound excessively general or even trivial, but it is certainly essential. In the absence of a time reference, the information that time coding can carry would be extremely limited. The aforementioned O'Keefe and Recce (1993) experiment is a good example for such a role of coherence (the theta rhythm). The fact that rate coding is common in sensory or motor neurons is also consistent with such a role of coherence: rate coding must be the only choice when a common time reference extending throughout the nervous system is not available. If the role of coherence in biological systems is to provide a time reference, we do not need to complicate the design of our network by adding a mechanism to produce coherence; we can provide a time reference from outside.

The BN is a variation of the integrate-and-fire neuron model (Glass & Mackey, 1979). We chose the integrate-and-fire neuron model because it is simple, but still can deal with time coding. Its internal potential increases in response to a postsynaptic current until it reaches a threshold level. When the potential touches the threshold level, it drops instantaneously to a relaxation level, producing a spike output. The behavior of an integrate-and-fire neuron appears to be monotonous: it repeats a regular cycle of charging and discharging. This unrealistically simple picture of an integrate-and-fire neuron is the result of an unrealistic assumption about the environment surrounding the neuron. A neuron in a biological neural network receives inputs from many different parts of the brain and is also subject to noise from inside and outside. Also, the same input can have different effects depending on the nature and the position of the synapse. Consideration of every detail of the interaction of a neuron with its environment would be impractical. Therefore, we made the following rather simplistic assumptions about the artificial environment surrounding an integrate-and-fire neuron. First, the environment provides a persistent *incoherent input* to the neuron, which helps the neuron keep active at an optimal operating point. Second, the environment also provides a persistent *coherent input* to the neuron, which is common to all the neurons in the same network and serves as a time reference.

It is possible to design a network such that it can provide itself with such incoherent and coherent inputs. For instance, some extra diffuse connection can be added between model neurons so that the neurons can feed each other with an incoherent input. Also, a part of the network can be used to build a pace maker that can provide a coherent input to other parts of the network. However, any attempt to model a biological neural network cannot avoid an approximation at some part of level of modeling. Here, we are approximating the mechanisms that are responsible for the incoherent and the coherent input by simply calling them 'the environment'. This way, we can focus on what happens when a time-coding-aware model network is subjected to such background activities. In this respect, a BN does not represent a neuron as it is, but a neuron augmented by incoherent activities and coherent activities in its environment. This is akin to the concept of an electron with an effective mass in solid state physics: an electron with an effective mass is not a model of an electron standing alone in the vacuum, but a model of an electron augmented by the influence of the atoms in the lattice surrounding it.

The following set of equations defines the BN:

$$\frac{dx_i}{dt} = c_i \tag{1}$$

$$\theta_i(t) = 1 \tag{2}$$

$$\rho_i(t) = \rho_0 \sin 2\pi f t \tag{3}$$

where $x_i(t)$, $\theta_i(t)$, and $\rho_i(t)$ are the internal potential, the threshold level, and the relaxation level of BN i, respectively. Eq. (1) defines the behavior of the internal potential $x_i(t)$ while it does not experience a relaxation. The internal potential $x_i(t)$ rises at a constant rate c_i, due to a constant feed by an incoherent source, until it reaches the threshold level $\theta_i(t)$. On touching the threshold level, the internal potential $x_i(t)$ drops instantly to the relaxation level $\rho_i(t)$. Eq. (2) defines the threshold level of BN i, which is constant when a BN has no connection from other BNs. Eq. (3) defines the behavior of the relaxation level, which maintains a sinusoidal oscillation of an amplitude ρ_0 and a frequency f due to a sinusoidal stimulation from a coherent source. Since the sinusoidal oscillation is extrinsic in nature, we call $\rho_i(t)$ a *driving signal* and, accordingly, ρ_0 and f a *driving amplitude* and a *driving frequency*, respectively. Also, we will omit the subscript of $\rho_i(t)$ hereafter since a driving signal is shared by all the BNs in a network.

Fig. 1a depicts the typical behavior of the BN: it repeats a

constant buildup and an instant relaxation while it is subjected to a driving signal $\rho(t)$. An important fact to emphasize again here is that a driving signal provides a time reference to the BN. This fact is more clearly demonstrated in Fig. 1b, that shows the firing times of the BN with respect to a modulus of 1, which is the period of the driving signal. Although the firing of the BN here is not completely phase locked to the driving signal, the firing time exhibits a clear structure when it is represented relative to the phase of the driving signal. In addition to providing a time reference, the driving signal turns the BN into a chaotic neuron. This fact is also demonstrated in Fig. 1b, where one can see that the firing-time pattern is bifurcating until it becomes chaotic as the amplitude of the driving signal increases. Such a dynamic behavior of a neuron subjected to an oscillatory input has been a subject of interest among many researchers (Aihara & Matsumoto, 1986; Hayashi, Ishizuka, Ohta & Hirakawa, 1982; Holden & Ramadan, 1981), and a detailed treatment of the same theme, applied to integrate-and-fire neurons, appears in Farhat and Eldefrawy (1991, 1992). Farhat and Eldefrawy discovered that an integrate-and-fire neuron can operate in dynamically distinct modes of operation depending on the amplitude and frequency of an applied oscillatory input, and showed that such a model

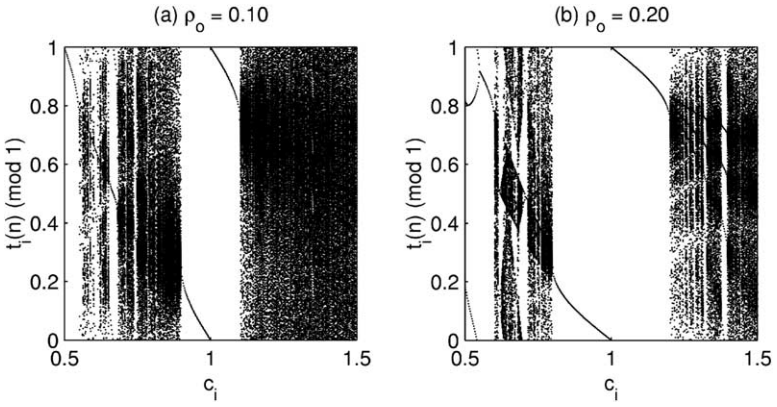

Fig. 2. The bifurcation diagrams of the firing time $t_i(n)$ (mod 1) of the BN as the bifurcation parameter c_i sweeps from 0.5 to 1.5: (a) $f = 1$ and $\rho_0 = 0.1$; (b) $f = 1$ and $\rho_0 = 0.2$.

page can be activated by the tuning of the driving frequency.[3]

The volume-holographic flavor of the BNN-2 implies that the memory capacity of the BNN-2 can be expanded by adding more recurring interconnections while the number of BNs is kept constant. This feature is in sharp contrast with that of traditional neural networks, where memory capacity is usually determined solely by the number of neurons. The volume-holographic flavor of the BNN-2 would not immediately translate into a practical advantage because adding more interconnections is often more costly than adding more neurons. Nevertheless, the volume-holographic flavor of the BNN-2 can explain one possible function of the dense, often recurring, connections observed in biological neural networks. The role of recurring connections in biological neural networks is difficult to understand from the viewpoint of traditional neural networks, e.g. the Hopfield networks, because, in these networks, adding recurring connections to an already fully-connected network is hardly meaningful. It should be emphasized that it is the time coding nature of the BNN-2 that enables its volume-holographic flavor.

In Section 2, we will explain the mechanism that makes the BN behave as an amplitude-to-phase converter. In Section 3, we will describe a pulse-coupling scheme that will make the BNN-2 an analog associative memory. In Section 4, we will suggest three different methods of training the BNN-2, thereby completing the definition of the BNN-2. In Section 5, we will test the proposed mechanism in a series of numerical simulations. In the final section, we will summarize our work presented in the paper and list some afterthoughts regarding the potential and the future of the BNN-2.

[3] The similarity between the role of the driving frequency in the memory recall of the BNN-2 and the role of the incidence angle of the refractive wave in volume holography will become more clear if one thinks of the reference beam in holography in terms of the angular spectrum of plane waves (Goodman, 1996).

2. Amplitude-to-phase transformation characteristics of the BN

In Fig. 1a, the ratio of the two perpendicular sides of the shaded triangle is equal to c_i, the linear buildup rate of the internal potential $x_i(t)$. Therefore, the n-th firing time of a BN $t_i(n)$ satisfies the following recursion when the BN is subjected to a sinusoidal oscillation given by Eq. (3):

$$t_i(n+1) = t_i(n) + \frac{1}{c_i} - \frac{\rho_0}{c_i}\sin 2\pi f t_i(n) \tag{4}$$

Fig. 2 shows the bifurcation diagram of the firing time $t_i(n)$ (mod 1) of a BN as the bifurcation parameter c_i ranges from 0.5 to 1.5. Note that these bifurcation diagrams look completely different from that shown in Fig. 1 since the buildup rate c_i is now chosen as a bifurcation parameter. In the period-1 windows of the bifurcation diagrams, which are centered around the point $c_i = 1$, the curves formed by fixed points are given by the following equation which can be derived from Eq. (4) by letting $t_i(n+1) = t_i(n) + 1/f$.

$$t_i^* = \lim_{n\to\infty} t_i(n) = \frac{1}{2\pi f}\arcsin\frac{1}{\rho_0}\left(1 - \frac{c_i}{f}\right) \tag{5}$$

If we let $c_i = f + u_i$, Eq. (5) becomes

$$t_i^* = -\frac{1}{2\pi f}\arcsin\frac{u_i}{f\rho_0} \tag{6}$$

When u_i is 0, t^*_i becomes 0. This means that the BN fires exactly at the beginning of the sinusoidal cycle of the driving signal $\rho(t)$. As u_i increases, however, a 'phase-lead' develops: the spiking of the BN starts to lead the beginning of the sinusoidal cycle. On the other hand, if u_i decreases, a 'phase-lag' develops: the spiking of the BN starts to lag behind the beginning of the sinusoidal cycle. According to Eq. (6), the phase of the spiking changes from $1/2\pi f$ to $-1/2\pi f$ when u_i changes from $-f\rho_0$ to $f\rho_0$. In other words, if we regard u_i as an external input, the BN is converting the external input to a phase shift in its output spike time.

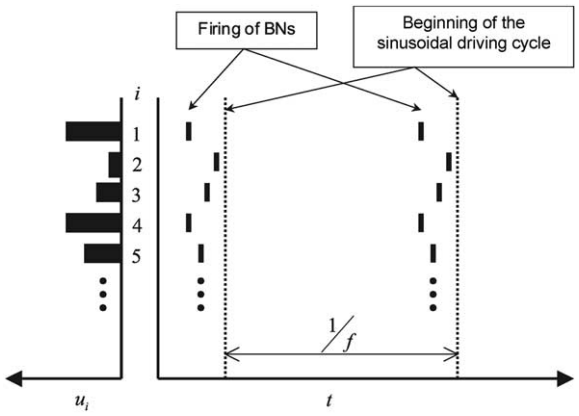

Fig. 3. Amplitude-to-phase transformation by a series of BNs.

Fig. 3 illustrates what the amplitude-to-phase conversion implies when it comes to a network of BNs. The bar graph on the left side shows an analog input pattern u_i. All the BNs in the network are phase locked to a driving signal and will respond to the analog input with proportional phase-leads in their firing times. Therefore, one can expect that the input pattern will be produced as a 'phase-lead pattern' in the firing times of the BNs as shown in the raster plot on the right side. This figure reminds us of a similar figure that appears in Hopfield (1995). He also suggested that a spiking neuron can convert an analog input to a phase shift when it is influenced by an oscillatory drive. It is interesting to see that we arrived at seemingly the same result as his, despite an apparent difference between his neuron model and the BN, our neuron model. As an aside, we would like to point out that this similarity will allow us to use his application ideas for the BN and its network. For instance, we will be able to use a BN network to detect an analog pattern, if we add a coherence-detector neuron which receives spikes from the BNs through properly time-delayed connections. However, the detailed conversion characteristics of the BN is different from that of his model neuron. For instance, the BN performs an almost linear conversion while his model performs a nonlinear conversion, e.g. a logarithmic conversion under a certain condition. Therefore, we would not be able to use the BN to explain the logarithmic intensity transformation of the visual system, for example. Note, however, that linear conversion is better than any other type of conversion, as far as an associative memory is concerned, since it can transfer information with the minimal loss in the presence of noise.

3. Pulse coupling between BNs

As shown in the previous section, a network of BNs, given an analog input pattern u_i, develops a phase-lead pattern which is related to the input pattern by Eq. (6). However, the induced phase-lead pattern will fade away as soon as the input is removed. It was speculated that the network may be able to maintain the induced phase-lead pattern even after the removal of the input if it can establish appropriate time-delayed interconnections between BNs. This idea is illustrated in Fig. 4 which shows a thought experiment that uses a minimal example network of two BNs. A step-by-step explanation of the experiment follows:

Fig. 4a. Initially, no input is provided to the BNs. According to Eq. (6), both BNs will be firing exactly at the beginning of the sinusoidal cycle of a driving signal. This means that they are in synchrony. However, this does not mean that they are phase locked to each other, but they are phase locked to the driving signal $\rho(t)$ which is common to all of them.

Fig. 4b. A positive external input is provided to BN 2. According to Eq. (6), BN 2 will develop a phase-lead of a small amount. This phase-lead will persist as long as the input is maintained.

Fig. 4c. Two time-delayed connections are added: one from BN 1 to BN 2 and the other from BN 2 to BN 1.

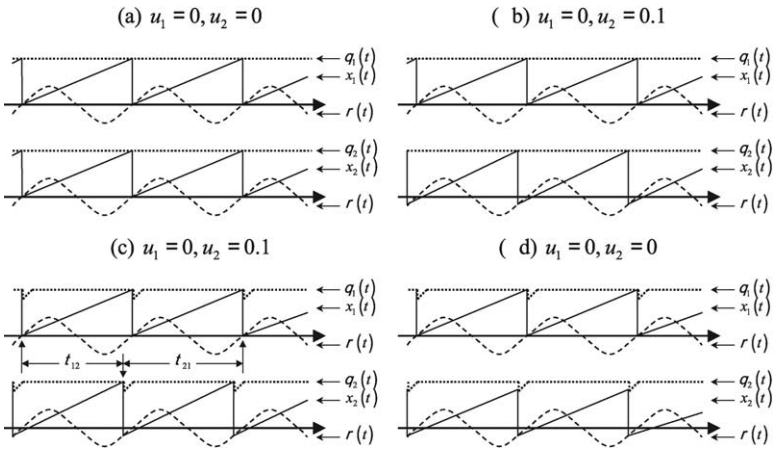

Fig. 4. A simple example to illustrate a possible mechanism to realize an associative memory using a BN network.

The lengths of the time delays are determined such that the situation may look as if the firings of the BNs are triggered by each other. The required lengths of the time delays, τ_{12} and τ_{21}, can be calculated by using Eq. (6).

Fig. 4d. The input to BN 2 is removed. What will happen then? We expect that the phase-lead of BN 2 which was induced by the input will now be maintained by the time-delayed connections.

The threshold coupling suggested in the thought experiment can be represented by the following equation:

$$\theta_i(t) = 1 - \delta_i(t) \tag{7}$$

where $\delta_i(t)$ denotes a perturbation in the threshold level induced by inputs from presynaptic BNs. Since the purpose of the threshold coupling is to maintain an induced phase-lead pattern, the coupling should have a short time constant so that it is capable of fine time resolution. A simple, but effective, type of coupling we chose is given by the following first order equation:

$$\frac{\mathrm{d}\delta_i}{\mathrm{d}t} = -\beta\delta_i + du_i^f(t) \tag{8}$$

or, equivalently,

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = -\beta(\theta_i - 1) - du_i^f(t) \tag{9}$$

where β is the reciprocal time constant of the restoring dynamics of the threshold level, $u_i^f(t)$ represents the network input, i.e. the input from presynaptic BNs, and d is a constant that controls the overall coupling strength of the BNs. The input $u_i^f(t)$ is the weighted sum of the delayed spike trains from the presynaptic BNs:

$$u_i^f(t) = \sum_{j=1}^{I}\sum_{k=1}^{K} w_{ij}^k y_j(t - \tau_{ij}^k) \tag{10}$$

where the variable $y_i(t)$ represents the output of BN i and can be approximated by a series of Dirac-delta functions, each of which represents a spike in the spike train:

$$y_i(t) = \sum_{n=1}^{N_i} \delta(t - t_i(n)) \tag{11}$$

where $t_i(n)$ is the n-th firing time of BN i, and N_i is the ordinal number of the latest firing of BN i. A synaptic connection from BN j to BN i is characterized by the two quantities, w_{ij}^k and τ_{ij}^k: the strength and time delay of the connection. In the current network design, the *weight matrix* w_{ij}^k is either 1 or 0, only indicating the existence of a connection, while the *delay matrix* τ_{ij}^k can take any real number centered around the average firing period of the BNs for the reason that will soon become clear. The superscript k signifies the possibility of recurring connections between a pair of BNs, and it, as it will turn out, is also the index of analog patterns to be remembered by the network. The required time delay τ_{ij}^k for the network to maintain a phase-lead pattern induced by an input u_i is given by Eq. (6), i.e.

$$\tau_{ij}^k = t_i^* - t_j^* + \frac{1}{f} = \frac{1}{f} - \frac{1}{2\pi f}\left(\arcsin\frac{u_i}{f\rho_0} - \arcsin\frac{u_j}{f\rho_0}\right) \tag{12}$$

where $1/f$ is added to $(t_i^* - t_J^*)$ in order to make sure that τ_{ij}^k is positive for all i, j, and k.

The thought experiment was about a network of two BNs which can store a single training pattern. Can the idea be extended to the case of a larger network and multiple training patterns? A series of simulations that follow will show that it is indeed the case, and that the associative memory becomes more robust as more BNs are involved in the network.

4. Training methods for BNN-2

Since the weight matrix w_{ij}^k is binary valued, it is more a design parameter that defines the structure of the BNN-2 than a target of training at least in this exploratory stage of the study of the BNN-2. Therefore, the following discussion of training methods concerns how to determine or adapt the delay matrix τ_{ij}^k given a set of training patterns. Consider the following two different choices of training methods:

Online training. Suppose that a BNN-2 already contains many recurring connections of a spread of time delays, but of negligibly small strengths. A training pattern, when applied to the network, will induce a phase-lead pattern in the firing times of the BNs. While the phase-lead pattern is maintained, we let each BN strengthen those incoming connections that deliver a spike from other BNs at the exact moment it fires. Repeat the same procedure for other training patterns.

Off-line training. Given a set of training patterns, we can calculate the corresponding time delays of connections by using Eq. (12). One can add all the required connections at once and in advance.

The online training would be a more biologically plausible way of training the BNN-2 since 'strengthen those incoming connections that deliver a spike from other BNs at the exact moment it fires' is stating exactly the principle of an online Hebbian learning (Hebb, 1949). Another advantage of the online training is that one does not need to depend on Eq. (12) which may become inaccurate if the driving signal of the BNN-2 is not perfectly sinusoidal. We are currently working on the online training of the BNN-2, but still have some practical problems to solve. In the following series of simulations, we will use a training method that we call 'quasi-online training'. It may be

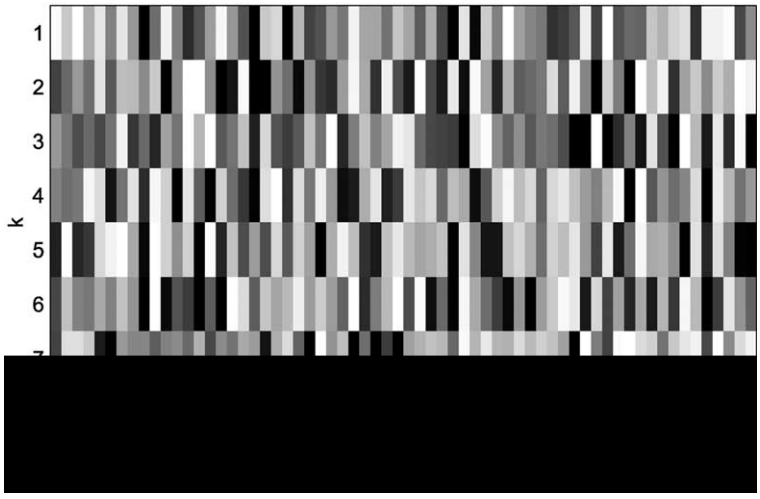

Fig. 5. Training patterns to be used in the simulations of the BNN-2 to test the associative memory function of the BNN-2. The intensity of the pixels in the plot denotes the values of ξ_i^k, uniform random numbers between 0 and 1; black for 0 and white for 1.

regarded as an online training method because it defines a simple action that can be done by individual BNs when a new training pattern is presented to a BNN-2. On the other hand, it is not an online training method in a true sence because it requires the addition (creation) of new connections to the network, which is not quite biologically plausible.

Quasi-online training. Suppose that BNN-2 contains no inter-BN connections initially. When a training pattern is applied to the network, it will induce a phase-lead pattern in the network. After the phase-lead pattern is stabilized, create time-delayed connections between BNs, where the amount of time delay of each connection is determined such that spikes from presynaptic BNs can arrive at postsynaptic BNs at the exact moment the postsynaptic BNs fire. Note that the computation of the time delays does not require the use of Eq. (12) since the time delay of a new connection is simply the difference in the firing times of a source and a destination BN. Repeat the same procedure for other training patterns.

5. Analog associative memory

The definition of the BNN-2 does not specify a particular form of network topology. BNs may be arranged in a one-dimensional or two-dimensional configuration, and they may have short range or long range interconnections. In the following numerical simulations of the BNN-2, we will assume that BNs form a two-dimensional lattice, and they have only local interconnections with neighboring BNs. Suppose that BNs are arranged in an $L \times L$ planar configuration ($I = L^2$). Each BN is connected to its neighbors that are within the Hemming distance r_c (inclusive), i.e.

$$w_{ij}^k = \begin{cases} 1 & d(i,j) \le r_c \\ 0 & d(i,j) > r_c \end{cases} \tag{13}$$

for all k, where $d(i,j)$ denotes the distance between BN i and BN j. In order to demonstrate the associative memory function of the BNN-2, we prepared eight training patterns ξ_i^k as shown in Fig. 5, where k is the index of training patterns, and i is the index of pixels. Since the pixel values of the training patterns range from 0 to 1, they need to be rescaled properly so that they may fall within the input dynamic range of the BNN-2, i.e.

$$u_i = \gamma f \rho_0 \xi_i^k \tag{14}$$

where $\gamma < 1$ is a positive scaling factor, and f and ρ_0 are the frequency and the amplitude of the driving signal. Eq. (12) now becomes

$$\tau_{ij}^k = t_i^* - t_j^* + \frac{1}{f} = \frac{1}{f} - \frac{1}{2\pi f}(\arcsin \gamma \xi_i^k - \arcsin \gamma \xi_j^k) \tag{15}$$

Preliminary numerical experiments showed that the network parameters, ρ_0, γ, and β, should be determined carefully for the optimal performance of the BNN-2 while the driving frequency f can be chosen to be 1, without loss of generality, by a proper normalization of the time variable. Currently, we are working hard toward the development of a systematic method for finding the optimal network parameters. Meanwhile, however, we will rely on the following rules of thumb, which we gathered from trials and errors.

- As ρ_0 becomes larger, the BNs tend to have a stronger synchronizing tendency, but seems to become less tolerant of an error in the time delays of the interconnections. A series of trials and errors led us to the conclusion that $\rho_0 = 0.1$ is a reasonable choice.

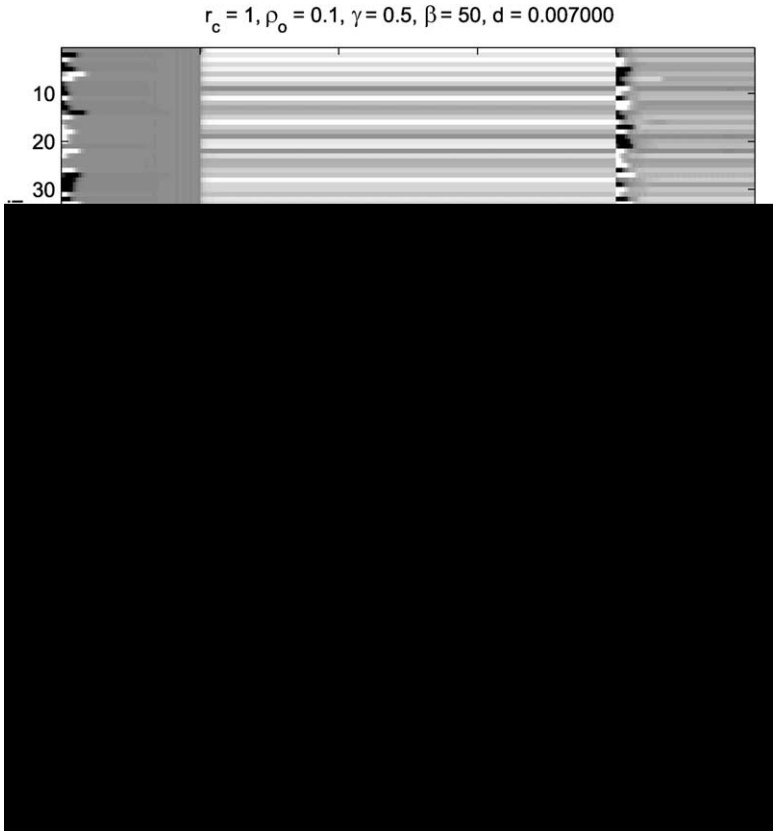

Fig. 7. Simulation 1: the time evolution patterns of the BNN-2 shown in a gray-scale raster plot, where black represents a firing-time lead of 0.05, and white represents a firing-time lag of 0.05: (top) for $d = 0.007$ and (bottom) for $d = 0.011$.

evolution is the same as that of the first experiment. The only difference is in the recall quality of the embedded pattern: the recalled pattern is in better agreement with the input pattern ξ_i^1 than that of the first case. Of course, this result does not mean that larger choice of d will always give better recall quality. If d is chosen too large, the behavior of the network becomes unpredictable.

5.2. Simulation 2: multiple attractors in the BNN-2

A similar network configuration as was used in the previous simulation is used to test if the BNN-2 can remember multiple patterns simultaneously. The first four patterns in Fig. 5 are used in this simulation. The required inter-BN connections can be added to the network in the same way as before except that the network now needs multiple, recurring connections between the same source and destination BNs. All the required connections are added to the network in advance according to the quasi-online training scheme defined in Section 4.

Suppose that the first training pattern ξ_i^1 is applied to the network as an input. The connection associated with this pattern will give rise to strong synchronized spikes in the threshold levels of the BNs. On the other hand, the other connections will result in a spread of small spikes. This effect is statistical in nature and, therefore, it is necessary to have large numbers of connections among BNs. In this simulation, the radius of connection r_c is chosen to be 3, which gives $M = 48$. The choice of the other network parameter values in this simulation is as follows: $f = 1, \rho_0 = 0.1, \gamma = 0.5, \beta = 200$, and $d = 0.0013$. This value of d is chosen as such based on the rough estimation given by Eq. (16): $d > 0.05/48 \approx 0.001$ for $M = 48$.

- At $t = 0$, the network is randomized, i.e. the internal potential of the BNs are initialized with random numbers.
- At $t = 50$, the first training pattern ξ_i^1 is applied to the network as an input, i.e. $u_i = \gamma f \rho_0 \xi_i^1$.
- At $t = 75$, the input is removed, i.e. $u_i = 0$.
- The above two steps are repeated for the remaining training patterns.
- At $t = 250$, a random pattern, which is not one of the four training patterns, is applied to the network as an input.
- At $t = 275$, the input is removed.

The simulation results are shown in Fig. 8 which contains two types of plots: a raster plot which was also used in the previous simulation, and a line plot, which shows the

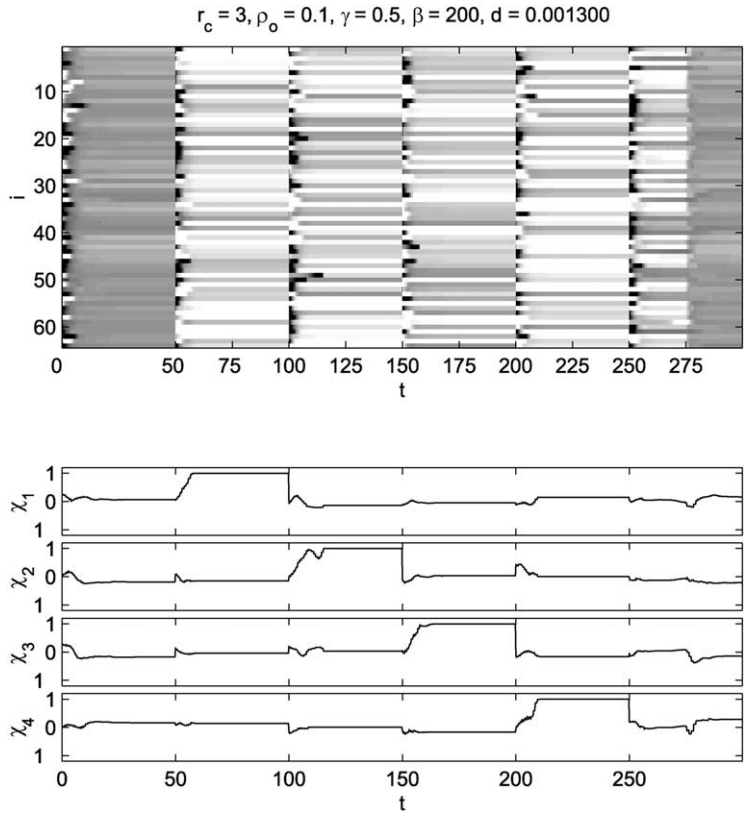

Fig. 8. Simulation 2: (top) the time evolution pattern of the BNN-2 shown in a gray-scale raster plot, where black represents a firing-time lead of 0.05, and white represents a firing-time lag of 0.05, and (bottom) the change of the correlation $\chi_k(t)$ in time.

temporal change of the correlation between the phase-lead pattern of the network and one of the four training patterns during the simulation:

$$\chi_k(t) = \frac{\sum_{i=1}^{I} (\xi_i^k - \bar{\xi}^k)(\zeta_i(t) - \bar{\zeta}(t))}{\sqrt{\sum_{i=1}^{I} (\xi_i^k - \bar{\xi}^k)^2 \sum_{i=1}^{I} (\zeta_i(t) - \bar{\zeta}(t))^2}} \tag{18}$$

where

$$\bar{\xi}^k = \frac{1}{I} \sum_{i=1}^{I} \xi_i^k \tag{19}$$

$$\bar{\zeta}(t) = \frac{1}{I} \sum_{i=1}^{I} \zeta_i(t) \tag{20}$$

and $\zeta_i(t)$ is defined in Eq. (17). According to this definition, $\chi_k(t)$ will take small value close to 0 when there is no correlation between ξ_i^k and $\zeta_i(t)$, but will be maximized ($\zeta_i(t) = 1$) when $\zeta_i(t)$ is proportional to ξ_i^k.

The simulation results demonstrate that there is a marked difference in the responses of the network to a known pattern and an unknown pattern. A phase-lead pattern induced by a known pattern persists after the removal of the input until a new input is applied. The response of the network is completely different for an unknown pattern: the network forgets the pattern as soon as the input is removed. The bottom line is that the BNN-2 is capable of remembering multiple analog patterns essentially without interference.

5.3. Simulation 3: large basins of attraction?

We have shown that the BNN-2 can maintain the phase-lead pattern which is induced by one of the training patterns (known patterns). On the other hand, it cannot maintain the phase-lead pattern which is induced by random patterns (unknown patterns). This demonstrates the basic memory capability of the BNN-2. However, this alone is not sufficient to make the BNN-2 a useful associative memory since an associative memory should be able to recognize a pattern which is known, but is distorted by noise. Also, the entire input pattern sometimes may not be available, and an associative memory is required to complete the entire pattern from a part of it. In terms of nonlinear dynamics, this problem reduces to that of the size of a basin of attraction. We have shown in the foregoing simulations that the training patterns form attractors in the BNN-2. However, we still need to examine how large are the basins of attraction associated with the attractors. In fact, we have already answered this question, though in part, in Simulation 1. The BNN-2, which was started with a random initial condition, could

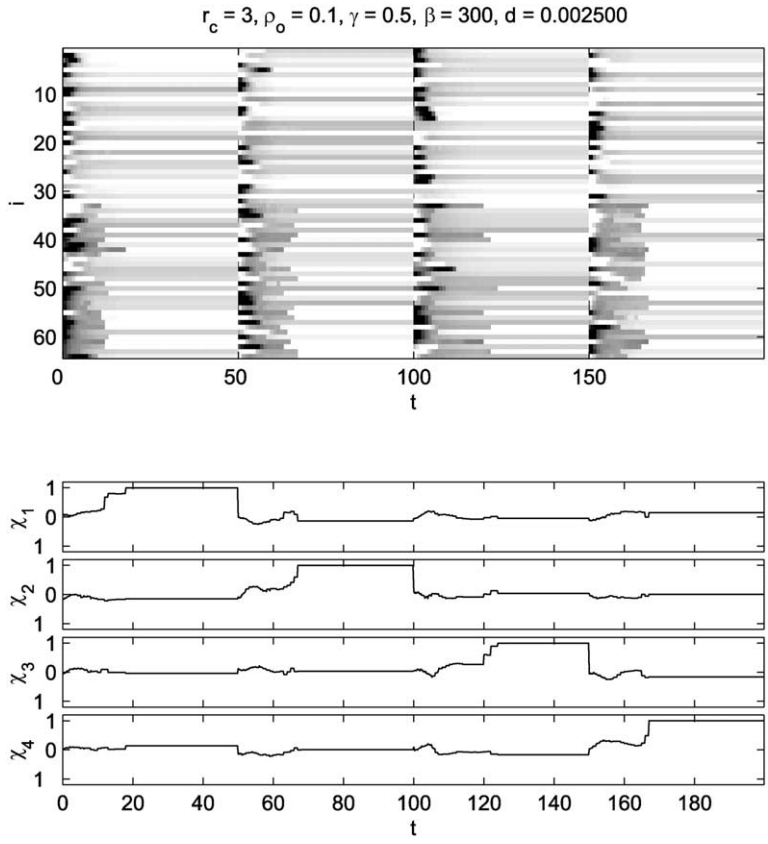

Fig. 9. Simulation 3 ($\beta = 300$): (top) the time evolution pattern of the BNN-2 shown in a gray-scale raster plot, where black represents a firing-time lead of 0.05, and white represents a firing-time lag of 0.05, and (bottom) the change of the correlation $\chi_k(t)$ in time.

successfully recall an embedded pattern. This means that the basins of attraction associated with the embedded pattern were occupying the whole input space. However, the answer is only to a special case of a single training pattern. Simulation 3 will answer the same question when BNN-2 is required to store more than one pattern. The choice of the network parameter values used in this simulation is summarized below: $r_c = 3$, $f = 1$, $\rho_0 = 0.1$, $\gamma = 0.5$, $\beta = 200$ and 300, and $d = 0.0025$. The same simulation will be repeated for the two different values of β.

We prepared corrupted versions of the first four training patterns, which were shown in Fig. 5, to be used in the simulation.

$$\tilde{\xi}_i^k \equiv \begin{cases} \xi_i^k & i \le I/2 \\ 0 & i > I/2 \end{cases} \tag{21}$$

The first half of each corrupted version is the same as that of the corresponding original pattern, and the second half is filled with zeros. All the connections with the required time delays to store the four *original* training patterns are established in advance according to the quasi-online training scheme defined in Section 4.

- At $t = 0$, the corrupted version $\tilde{\xi}_i^1$ of the first training pattern is applied to the network as an input, i.e. $u_i = \gamma f \rho_0 \tilde{\xi}_i^1$. Then, the network is allowed to run until $t = 50$ so that it can reconstruct the whole pattern.
- The above step is repeated for the other corrupted training patterns: $\tilde{\xi}_i^2$ at $t = 50$, $\tilde{\xi}_i^3$ at $t = 100$, and $\tilde{\xi}_i^4$ at $t = 150$.

The result of the case of $\beta = 300$ is shown in Fig. 9. We can see an apparent visual difference in the pattern of the network response to the inputs in the first and the second halves of the network. A longer transient period in the second half of the network indicates the network's effort to reconstruct the missing part in the input patterns. In the case of $\beta = 200$, the network could reconstruct the missing part successfully for the first, third and fourth training patterns, but not for the second pattern. It seemed that the network tends to perform better when the parameter β is larger. However, the network model will depart further from biological reality if the parameter β becomes too large.

5.4. Simulation 4: volume holographic memory

According to Eq. (15), the time delays associated with the connections required to store a pattern in the BNN-2 are dependent on the driving frequency f. This means that the BNN-2 can recall a stored pattern only when the same driving signal that was present at the time of training is present.

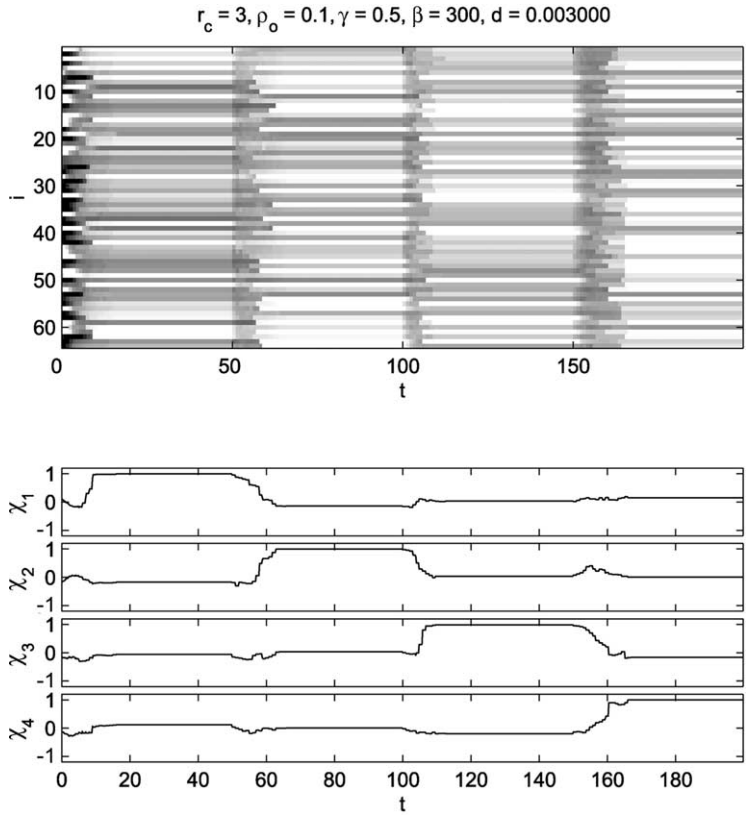

Fig. 10. The first run of Simulation 4: (top) the time evolution pattern of the BNN-2 shown in a gray-scale raster plot, where black represents a firing-time lead of 0.05, and white represents a firing-time lag of 0.05, and (bottom) the change of the correlation $\chi_k(t)$ in time.

This situation is reminiscent of volume holography. For a volume hologram to reconstruct a recorded image, it is necessary that a reference beam should be supplied exactly at the same angle as it was in the recording step. In fact, this sensitivity to the angle of the reference beam enables volume holography to store multiple patterns in a single photorefractive crystal. It appears that the driving frequency f of the BNN-2 plays the role of the angle of a reference beam in volume holography. Therefore, the question which we are about to answer in the simulation is this: can the BNN-2 store different patterns at different driving frequencies without causing interference among the patterns?

The values of the network parameters used in this simulation are as follows: $r_c = 3$, $f = 1$, $\rho_0 = 0.1$, $\gamma = 0.5$, $\beta = 300$, and $d = 0.003$. The same simulation was repeated twice for different numbers of training patterns.

The result of the first run is shown in Fig. 10. The time-delayed connections for the first four training patterns shown in Fig. 5 are added to the network according to the quasi-online training scheme. In doing so, we used different driving frequencies for different training patterns: $f = 1$, 1.02, 1.04 and 1.06 for the four training patterns, respectively. At $t = 0$, the network is randomized, i.e. the internal potentials of the BNs are initialized with random numbers. While the simulation is running, the frequency of the driving signal is changed discontinuously as follows: $f = 1$ at $t = 0$, $f = 1.02$ at $t = 50$, $f = 1.04$ at $t = 100$, and $f = 1.06$ at $t = 150$. Each of the four embedded patterns shows up when the driving frequency is switched to the frequency which was used in the training of the pattern. This experimental result clearly demonstrates that the memory of the BNN-2 is, indeed, dependent on the driving frequency. In other words, it demonstrates the multiplexing nature of the BNN-2 associative memory.

In the second run, we used all the eight training patterns shown in Fig. 5. The first four and the second four training patterns were learned by the network under the driving frequencies of $f = 1$ and 1.02, respectively. We designed this experiment in order to check if the two memory spaces associated with the two different driving frequencies are independent and do not interfere. The simulation procedure is similar to that of Simulation 2 except that the fifth training pattern is used in place of the random pattern which was used in Simulation 2. The simulation result is shown in Fig. 11.

- At $t = 0$, the network is randomized, i.e. the internal potential of the BNs are initialized with random numbers.
- At $t = 50$, the first training pattern ξ_i^1 is applied to the network as an input, i.e. $u_i = \gamma f \rho_0 \xi_i^1$.
- At $t = 75$, the input is removed, i.e. $u_i = 0$.
- The above two steps are repeated for the other three training patterns.
- At $t = 250$, the fifth pattern, which is not one of the four

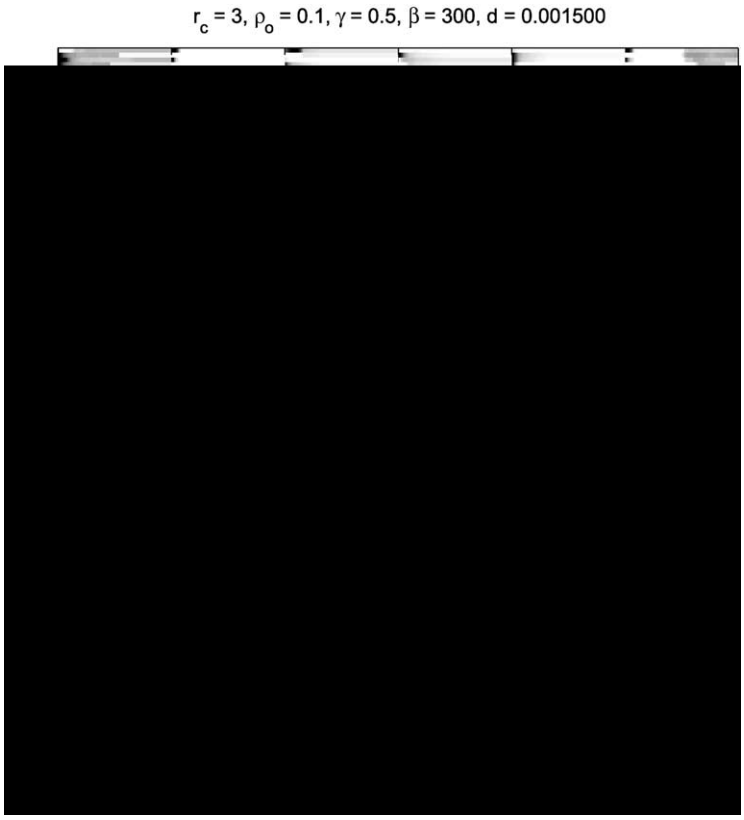

Fig. 11. The second run of Simulation 4: (top) the time evolution pattern of the BNN-2 shown in a gray-scale raster plot, where black represents a firing-time lead of 0.05, and white represents a firing-time lag of 0.05, and (bottom) the change of the correlation $\chi_k(t)$ in time.

training patterns which were learned by the network under the driving frequency of 1, is applied to the network as an input.
- At $t = 275$, the input is removed.

Throughout the simulation, the driving frequency was kept constant at $f = 1$. Up to our expectation, the network cannot recognize the fifth pattern which was learned under a different driving frequency ($f = 1.02$). The network is treating the fifth pattern in the same way as it did a random pattern in Simulation 2. This experiment result clearly demonstrates the independence of the memory spaces associated with different driving frequencies.

6. Conclusions

We applied the amplitude-to-phase converting function of the BN to the design of an analog associative memory, whose principle was illustrated using a simple thought experiment. The idea was verified through a series of numerical simulations. While the first three simulations demonstrated the basic characteristics of the BNN-2 as an analog associative memory, the last simulation emphasized the multiplexing, or 'volume-halographic' flavor of the BNN-2 memory. We are particularly interested in the following characteristics of the BNN-2.

1. The memory of the BNN-2 is dependent on the presence of a driving signal of a specific frequency. This implies the following two consequences. First, the BNN-2 can be *turned on or off* by an external control of a driving signal. In the absence of a relevant driving signal, the response of the BNN-2 is largely noisy and meaningless. It becomes active only when a specific driving signal is provided. Here, the role of a driving signal strongly reminds us of the function of the thalamic reticular complex suggested by Crick in his searchlight hypothesis (Crick, 1984; Crick & Koch, 1990). Second, the memory of the BNN-2 is driving-frequency dependent, i.e. it is context dependent. The driving frequency may be used as a context key to access different pages in the BNN-2 memory. This also means that the same neural network can represent a completely different memory after a kind of mode switching. We are almost tempted to translate the context-dependency of the BNN-2 into the language of neuroscience, but will postpone it until we become more confident in the BNN-2 as a model of a biological neural network.
2. The BNN-2 utilizes multiple recurring connections between the same source and destination pair of BNs. Considering the extremely high connectivity of the brain (of the order of 10^4 for each neuron), the dramatic reduction to a single connection between a source and a

destination, as seen in most of the network models, should be an oversimplification. Such a reduction may be justified under a rate-coding approximation, but may not if the individual timing of the neuronal spikes are to be counted. With such multiple recurring connections, a neuron now can signal to the destination neurons with a temporal pattern, not merely with a single spike. The BNN-2 provides an example how such recurring connections in the brain can be utilized.

The amplitude-to-phase converting function of the BN alone may be able to find many useful applications. It can be used to convert a rate-coded signal to a time-coded signal which may be more appropriate as an input to another spiking neuron network. It can be also useful in building a pattern matching network similar to the one suggested by Hopfield (1995) or building a radial basis function network similar to those that appeared in Natschlaeger and Ruf (1998).

As a final note, we have to mention that the work presented in this paper is still exploratory in nature, and our exploration of the possibility of the BNN-2 is by no means complete. At the time of writing, we are diligently working toward the establishment of a more theoretical ground for the BNN-2. A few immediate goals of our current study of the BNN-2 include the development of a biologically-plausible online training method, the estimation of the memory capacity of the BNN-2, and the development of a systematic method for the optimal choice of the network parameters.

Acknowledgements

This research was supported by the Office of Naval Research under grant no. N00014-01-1-0418.

References

Adrian, E. D. (1926). The impulses produced by sensory nerve endings. *Journal of Physiology*, *61*, 49–72.

Aihara, K., & Matsumoto, G. (1986). Chaotic oscillations and bifurcations in squid giant axons. In A. V. Holden, *Chaos* (pp. 257–269). Princeton University Press.

Basar, E. (1983). Synergetics of neural populations: survey on experiments. In E. Basar, H. Flohr, H. Haken & A. J. Mandell, *Synergetics of the brain* (pp. 183–200). Springer.

Bibbig, A., Wennekers, T., & Palm, G. (1995). A neural network model of the cortico-hippocampal interplay and the representation of contexts. *Behavorial Brain Research*, *66*, 169–175.

Bressler, S. L., & Freeman, W. J. (1980). Frequency analysis of olfactory system EEG in cat, rabbit, and rat. *Electroencephalography and Clinical Neurophysiology*, *50*, 19–24.

Caianiello, E. R. (1961). Outline of a theory of thought and thinking machines. *Journal of Theoretical Biology*, *1*, 204–235.

Cragg, B. G., & Temperley, H. N. V. (1955). Memory: the analogy with ferromagnetic hysteresis. *Brain*, *78* (II), 304–316.

Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences of the USA*, *81* (14), 4586–4590.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? *Biological Cybernetics*, *60*, 121–130.

Farhat, N. H., & Eldefrawy, M. (1991). The bifurcating neuron. In *Digest Annual OSA Meeting*, San Jose, CA (p. 10).

Farhat, N. H., & Eldefrawy, M. (1992). The bifurcating neuron: characterization and dynamics. In *Photonics for computers, neural networks, and memories*, vol. 1773, SPIE Proceedings (pp. 23–34), San Diego, CA, July, 1992. SPIE.

Freeman, W., & Skarda, C. A. (1985). Spatial EEG patterns, nonlinear dynamics and perception: neo-Sherringtonian view. *Brain Research Reviews*, *10*, 147–175.

Freeman, W., & van Dijk, B. W. (1987). Spatial patterns of visual cortical fast EEG during conditioned reflex in a rhesus monkey. *Brain Research*, *422*, 267–276.

Fukai, T., & Shiino, M. (1995). Memory recall by quasi-fixed-point attractors in oscillator neural networks. *Neural Computation*, *7*, 529–548.

Gerstner, W., & van Hemmen, J. L. (1992). Associative memory in a network of spiking neurons. *Network*, *3*, 139–164.

Glass, L., & Mackey, M. C. (1979). A simple model for phase locking of biological oscillators. *Journal of Mathematical Biology*, *7*, 339–352.

Goodman, J. (1996). *Introduction to Fourier optics*, New York: McGraw-Hill.

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences of the USA*, *86*, 1698–1702.

Hayashi, H., Ishizuka, S., Ohta, M., & Hirakawa, K. (1982). Chaotic behaviour in the onchidium giant neuron under sinusoidal stimulation. *Physics Letters A*, *88* (8), 435–438.

Hebb, D. O. (1949). *The organization of behavior: a neuropsychological theory*, New York: Wiley.

Herz, A. V., Li, Z., & van Hemmen, J. L. (1991). Statistical mechanics of temporal association in neural networks with transmission delay. *Physical Review Letters*, *66* (10), 1370–1373.

Hilborn, R. C. (1994). *Chaos and nonlinear dynamics, an introduction for scientists and engineers*, New York: Oxford University Press.

Holden, A. V., & Ramadan, S. M. (1981). The response of a molluscan neuron to a cyclic input: entrainment and phase-locking. *Biological Cybernetics*, *41*, 157–163.

Holden, A. V., Hyde, J., Muhamad, M. A., & Zhang, H. G. (1992). Bifurcating neuron. In J. G. Taylor & C. L. T. Mannion, *Coupled oscillating neurons* (pp. 41–80). London: Springer-Verlag.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, *79*, 2554–2558.

Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, *376* (6), 33–36.

Krieger, D., & Dillbeck, M. (1987). High frequency scalp potentials evoked by a reaction time task. *Electroencephalography and Clinical Neurophysiology*, *67*, 222–230.

Kuffler, S. W., Nicholls, J. G., & Martin, A. R. (1992). *From neuron to brain* (3rd ed.). Sinauer Associates Inc.

Little, W. A., & Shaw, G. L. (1975). A statistical theory of short and long term memory. *Behavioral Biology*, *14*, 115–133.

Maass, W., & Natschlaeger, T. (1998). Associative memory with networks of spiking neurons in temporal coding. In L. S. Smith & A. Hamilton, *Neuromorphic systems: engineering silicon from neurobiology* (pp. 21–32). Singapore: World Scientific.

Milner, P. M. (1974). A model for visual shape recognition. *Psychology Review*, *81*, 521–535.

Natschlaeger, T., & Ruf, B. (1998). Spatial temporal pattern analysis via spiking neurons. *Network: Computation in Neural Systems*, *9*, 319–332.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map:

preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34* (1), 171–175.

O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3* (3), 317–330.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1996). *Spikes—exploring the neural code*, Cambridge, MA: MIT Press.

von der Malsburg, C. (1981). The correlation theory of brain function. In E. Domany, J. L. van Hemmen & K. Schulten (Eds.), *Models of neural networks II*. Springer-Verlag. Internal Report 81-2. Gottingen: Max-Planck Institute for Biophysical Chemistry.

Wennekers, T., Sommer, F. T., & Palm, G. (1995). Interactive retrieval in associative memories by threshold control of different neural models. In H. J. Herrmann, D. E. Wolf & E. Poeppel, *Supercomputing in brain research: from tomography to neural networks* (pp. 301–319). Singapore: World Scientific.