5-2013

Behind the magical numbers: Hierarchical chunking and the human working memory capacity

Guoqi LI

Ning NING

Kiruthika RAMANATHAN
Singapore Management University, kiruthikar@smu.edu.sg

Wei HE

Li PAN

See next page for additional authors

Behind the magical numbers: Hierarchical chunking and the human working memory capacity

Article *in* International Journal of Neural Systems · August 2013

DOI: 10.1142/S0129065713500196 · Source: PubMed

6 authors, including:

Guoqi Li
Tsinghua University
215 PUBLICATIONS **4,102** CITATIONS
SEE PROFILE

Ning Ning
Tencent
27 PUBLICATIONS **337** CITATIONS
SEE PROFILE

Kiruthika Ramanathan
Agency for Science, Technology and Research (A*STAR)
81 PUBLICATIONS **586** CITATIONS
SEE PROFILE

L.P. Shi
Agency for Science, Technology and Research (A*STAR)
132 PUBLICATIONS **3,616** CITATIONS
SEE PROFILE

International Journal of Neural Systems, Vol. 23, No. 4 (2013) 1350019 (12 pages)

DOI: 10.1142/S0129065713500196

World Scientific
www.worldscientific.com

BEHIND THE MAGICAL NUMBERS: HIERARCHICAL CHUNKING AND THE HUMAN WORKING MEMORY CAPACITY

GUOQI LI*, NING NING, KIRUTHIKA RAMANATHAN, WEI HE and LI PAN
Data Storage Institute, Agency for Science
Technology and Research, A STAR*
5 Engineering Drive 1, Singapore 117608, Singapore
LI_Guoqi@dsi.a-star.edu.sg

LUPING SHI
Department of Precision Instrument
Tsinghua University, Beijing 100084, P. R. China

Accepted 24 April 2013
Published Online

To explore the influence of chunking on the capacity limits of working memory, a model for chunking in sequential working memory is proposed, using hierarchical bidirectional inhibition-connected neural networks with winnerless competition. With the assumption of the existence of an upper bound to the inhibitory weights in neurobiological networks, it is shown that chunking increases the number of memorized items in working memory from the "magical number 7" to 16 items. The optimal number of chunks and the number of the memorized items in each chunk are the "magical number 4".

Keywords: Chunking; working memory; hierarchical structure; capacity; magical numbers.

1. Introduction

It is well known that working memory (WM)[1,2] has a limited capacity.[3] Miller[4] summarized evidence that people can remember about "magical number 7" items in WM tasks. However, recent empirical evidence has led psychologists to propose that this number is more of a rhetorical device than a real capacity limit, and to conclude that the capacity limit of WM averages around "magical number 4",[5] although the real upper limit is still under debate.[6]

On the other hand, psychologists agree that chunking, the process by which individual units of information are grouped into clusters, increases the capacity limit of WM.[7,8] This makes chunking a useful tool for memorizing large amounts of information. For example, a phone number 93208215 may be chunked into 9320-8215. Combining disparate individual elements into large blocks facilitates information storage and recall.

Theoretical analysis of the capacity limit of WM due to chunking is an interesting and open problem.[9,10] Literature studies show that there are very few works focusing on the theoretical analysis of the working memory capacity. A recent work by Bick and Rabinovich[11] modeled sequential WM based upon winnerless competition (WLC)[12] among representations of items and showed that the upper bound of WM is consistent with the "magical number 7".

In this paper, based on the Rabinovich's work, we propose a hierarchical sequential working memory (HSWM) model and use it to analyze theoretically the chunking of sequential information. According to Cowan's work, we considered WM as a set of activations within long-term memory.[13] In this case, chunking with a WM context could be possible. We further relate our analysis to the relative lateral inhibition in a neuronal network. Chunks in WM are modeled as a two-layered hierarchical bidirectional

inhibition connected neural network. The model consists of a single activating parent network (PN) connected to several activated child networks (CN). Each neuron in the PN is connected to a specific CN respectively and together, form one chunk. The WLC between neurons in both PN and CN is imitated by generalized Lotka–Volterra equations.[12] A winning neuron in the PN activates its connected CN. When a recall cue is given, the model presents a trace containing temporary winner neurons among the multiple chunks. The trace reflects the sequential memory recall. The model analysis and simulations suggest what chunking changes the capacity limit of human WM from seven items to four chunks, with each chunk consisting of four items (brings the total capacity to 16 items). On the other hand, we agree that normal human beings may be unable to recall sequences of 16 elements in working memory. 16 items could be the upper bound of WM for most persons. As shown in Fig. 5 later, the capacity of WM is between 12–16 items through chunking. So our conclusion is that the upper bound of WM is four chunks and four items in each chunk. Actually the items in each chunk is not necessary to be four items. Two or three items are also possible.

We would like to reiterate that the model proposed here is a theoretical analysis which gives the upper bound capacity of working memory, it is considered as a set of activations within long-term memory.[13] However, we understand that it is difficult to analyze working memory together with long term memory, especially in terms of memory capacity, as how to separate the two is still mysterious. In this paper, the hierarchical structure of the proposed HSWM is considered as the long term memory. The encoding of working memory can then be regarded as the neural bias and the inhibitory weights in HSWM. In analyzing the WM capacity, we mainly focus on the relative inhibition index of the encoded HSWM, which enables us to analyze the WM capacity without considering how long term memory involves in. In fact, the only paper we have found, from which we could infer such an upper bound is by Ref. 14. The experiments were conducted on amnesiac patients and the dependency of the chunking process on long term memory could reasonably be ignored. The results of their experiments show that amnesiac patients could recall a maximum of 15 items in the short term. Based on current literature, we think that this offers sufficient experimental justification for our model.

The significance of this work is that the proposed HSWM model can explain how chunking influences the upper bound of WM. While this bound depends on the inhibitory connections in the network, we will show that it can be extended to 16 items through chunking and derive that the optimal number of chunks and the items of each chunk are both four. This is notable, as it shows the relationship between the boundary capacity of WM and the magical numbers 7 and 4, which are points of debate in the psychology of WM.

2. Sequential Working Memory Model

WM is known as short-term memory[15] that acts as a kind of scratchpad for temporary recall of the information. There are works use the "attractors",[16–20] which involves the points or sets that get close enough to the attractor remain close even if slightly disturbed in attractor networks, to model WM. However, different from attractors, the dynamic in WM may not stay at any temporary recalled information. So winnerless competition (WLC) between neural representations is hypothesized to be the main mechanism for memory retrieval in WM.

On the other hand, the generalized Lotka–Volterra equations are a set of equations more general than either the competitive or predator–prey examples of Lotka–Volterra equations, which can be used to model competition between an arbitrary number of species. These species usually form an inhibited connected network. This is the reason why the generalized Lotka–Volterra equations can be used to describe the dynamical process of WLC in an inhibition connected network,[12] which is the basis of the HSWM described later.

As seen HSWM later, the PN neurons govern the activation of the CN networks. All these sub networks can be described by the same WLC process with different time constants. Usually the transition between neural activations in the PN is slower than that in the CN. We therefore modify the generalized Lotka–Volterra model in Ref. 11 by adding a time constant $\tau > 0$ as below:

$$\dot{x}_i = \tau \cdot x_i \cdot (\sigma_i - x_i + \Sigma_{j \neq i} \rho_{ij} x_j) + v_i \qquad (1)$$

for $i, j = 1, \ldots, n$, where n is the number of neurons in the neural network, τ reflects the rate of activation and decay of neurons, $x_i \geq 0$ is the output neural activity of neuron i, σ_i is the positive neural bias, $\rho_{ij} \leq 0$ for $i \neq j$ is the inhibitory weight from the neuron j to neuron i, and v_i is the external noise in the interval $(0, \varepsilon]$ where ε is a small positive constant. Note that $\rho_{ij} \neq \rho_{ji} \leq 0$ in this bidirectional inhibition connected network, and σ_i and $\rho_{ij} \leq 0$ for all i and j are considered as functions of the input stimulus.

Remark 1. In (1), each neuron represents one item, for example, one digital number. The neural activities, whose values are variable and described by WLC process by using generalized Lotka–Volterra equations, basically represent the level of activity of each neuron in a spiking neural network.[21–31] At any given time, the neuron that possesses the maximum neuron activity value is the temporal winner. And the items that neuron represents will be recalled. The model we proposed may be one possible way to describe the WM mechanism and definitely there are many other ways. For example, the mean-field approach[32–34] is a very powerful and influential way. ⋄

Let $\mathbf{x} = [x_1 \cdots x_n]^T$. We rewrite the equations in (1) as

$$\dot{\mathbf{x}} = \tau \cdot \frac{dF(\mathbf{x})}{d\mathbf{x}} + \mathbf{v}, \tag{2}$$

where

$$\begin{aligned} \frac{dF(\mathbf{x})}{d\mathbf{x}} &= \left[\frac{df_1(\mathbf{x})}{d\mathbf{x}} \cdots \frac{df_i(\mathbf{x})}{d\mathbf{x}} \cdots \frac{df_n(\mathbf{x})}{d\mathbf{x}}\right]^T \\ \frac{df_i(\mathbf{x})}{d\mathbf{x}} &= x_i \cdot (\sigma_i - x_i + \Sigma_{j \neq i} \rho_{ij} x_j) \end{aligned} \tag{3}$$

and $\mathbf{v} = [v_1 \cdots v_n]^T$. A temporal winner in the dynamical WLC is the one which preserves the maximum activity in $\mathbf{x}$ at a given time. We divide the n neurons into two sets. The first set I has $\mathbf{k}$ neurons which are likely to be winners (connection-friendly) sometime in the WLC. The other set J contains $n - \mathbf{k}$ neurons which are unlikely to be winners (connection-unfriendly) in the process. Let

$$I = \{I_1, \ldots, I_{\mathbf{k}}\} \tag{4}$$

be the neuron indices of the first set. The other set is

$$J = \{J_1, \ldots, J_{n-\mathbf{k}}\}. \tag{5}$$

As seen in (8)–(10) later, some weights $-\rho_{ij}$ between the neurons in the set I (belonging to $S_{k,\mu}$) are much smaller that other weights between the neurons in the set J (belonging to $\tilde{S}_{k,\mu}$). This is why the set I is called connection-friendly and the set J is called connection-unfriendly.

Now we investigate the design of σ_i and ρ_{ij} for all i and j for the network to exhibit its memory trace. Assume that the trace of temporal winners to encode into (1) is given as

$$I_k \to I_{k+1} \to \cdots \to I_{\mathbf{k}}, \tag{6}$$

where $1 \leq k \leq \mathbf{k}$ in set I. Generally speaking, the support of longer sequences requires increasing excitation, $\{\sigma_{I_k}\}_{k=1}^{\mathbf{k}}$ for the neurons in I is required to have a geometric growth. Let $\{\sigma_{I_k}\}_{k=1}^{\mathbf{k}}$ be a sequence of σ_{I_k} with k changing from 1 to $\mathbf{k}$. As in Ref. 11 the Fibonacci sequence is used to encode values for $\{\sigma_{I_k}\}_{k=1}^{\mathbf{k}}$ such that

$$\sigma_{I_1} = \sigma_0, \quad \sigma_{I_2} = \mu\sigma_{I_1}, \quad \text{for } \sigma_0 > 0, \ \mu > \frac{1}{2} \tag{7}$$

and $\sigma_{I_{k+1}} = \sigma_{I_k} + \sigma_{I_{k-1}}$ when $k \geq 2$, where σ_0 is a random value uniformly distributed over the interval $(0, 1]$. The minus of the inhibition weight $-\rho_{ij}$ is designed as a random variable belonging to either $S_{\mathbf{k},\mu}$ or $\tilde{S}_{\mathbf{k},\mu}$ as follows:

$$-\rho_{ij} \in \begin{cases} S_{\mathbf{k},\mu} & i, j \in I, \text{ and are adjacent in } I. \\ \tilde{S}_{\mathbf{k},\mu} & \text{otherwise.} \end{cases} \tag{8}$$

$$S_{\mathbf{k},\mu} = \left(\left(\max_{k \in \{1,\ldots,\mathbf{k}-1\}} \frac{\sigma_{I_{k+1}}}{\sigma_{I_k}} - \frac{1}{2}\right), \min_{k \in \{1,\ldots,\mathbf{k}-\mathbf{1}\}} \frac{\sigma_{I_{k+1}}}{\sigma_{I_k}}\right), \tag{9}$$

$$\tilde{S}_{\mathbf{k},\mu} = \left(\left(\max_{k,k' \in \{1,\ldots,\mathbf{k}\}} \frac{\sigma_{I_k}}{\sigma_{I_{k'}}} + 1\right), +\infty\right). \tag{10}$$

The domain of μ in (7) is

$$\mu \in \bar{\mu} = \left\{\mu > \frac{1}{2} \text{ and } S_{\mathbf{k},\mu} \neq \emptyset\right\}. \tag{11}$$

It is well-known that, the value of $\{\sigma_{I_{k+1}}/\sigma_{I_k}\}$ in a Fibonacci sequence approaches the "golden ratio" $g = \frac{\sqrt{5}+1}{2}$. The first three terms of $\{\sigma_{I_k}\}$ are $\sigma_0, \mu\sigma_0, (\mu+1)\sigma_0$. If $S_{\mathbf{k},\mu} \neq \emptyset$, we have $\frac{\mu+1}{\mu} - \frac{1}{2} < \mu$

and $\mu - \frac{1}{2} < \frac{\mu+1}{\mu}$. It can be solved that

$$\mu \in \bar{\mu} = \left(\frac{\sqrt{17}+1}{4}, 2\right). \tag{12}$$

Exploring the property of the Fibonacci sequence, we have $\sigma_{I_k} = (a_0 g^k - b_0(1-g)^k)\sigma_0$ where a_0 and b_0 are the parameters such that $\sigma_{I_1} = \sigma_0$ and $\sigma_{I_2} = \mu\sigma_0$.

Remark 2. Once $-\rho_{ij}$ is assigned to an interval $S_{k,\mu}$ or $\tilde{S}_{k,\mu}$, a value will be chosen uniform randomly in that interval. Note that both the $S_{k,\mu}$ and $\tilde{S}_{k,\mu}$ change with the length of the sequence (denoted as $\mathbf{k}$) in set I as shown in (8). When an incoming input is encoded as part of a chunk, the values of $-\rho_{ij}$ are chosen based on Eqs. (7) to (11). Once the values are encoded, they are not subsequently changed during the recall process. Therefore, once encoding is complete, the synaptic connections are independent of neural activity, but influence the activity of the neurons, as described in Eq. (1). $\diamond$

Theorem 1. *For system* (1) *in the presence of a stimulus with sufficiently small noise, the encoding process* (7)–(10) *guarantees that the trace of temporal winner neurons is* $\{I_k \to I_{k+1} \to \cdots \to I_{\mathbf{k}}\}$ *provided that the initial activation output state* $\mathrm{x} = A_{I_k} = [0, \ldots, \sigma_{I_k}, \ldots, 0]$ *where* $I_k \in I$ *with* σ_{I_k} *is the* I_k*th entry.*

A mathematical proof of the theorem is given in Appendix. Now we define the relative lateral inhibition index of the neural circuit (1) as

$$\varphi(\mathbf{k}, \mu) = \frac{\inf(\tilde{S}_{\mathbf{k},\mu})}{\sup(S_{\mathbf{k},\mu})}, \quad \mu \in \bar{\mu}, \tag{13}$$

which reflects the need of excitation to support the encoding procedure. A higher relative lateral inhibition index implies that a higher excitation is required to encode information in the circuit. It is therefore necessary to minimize $\varphi(\mathbf{k}, \mu)$ with respect to $\mu \in \bar{\mu}$. Based on the analysis of the Fibonacci sequence, we have

$$\inf(\tilde{S}_{\mathbf{k},\mu}) = \frac{\sigma_{I_{\mathbf{k}}}}{\sigma_0} + 1 = a_0 g^{\mathbf{k}} - b_0(1-g)^{\mathbf{k}} + 1$$

$$\sup(S_{\mathbf{k},\mu}) = \min\left\{\mu, 1 + \frac{1}{\mu}, 1 + \frac{1}{1+\frac{1}{\mu}}, \ldots\right\}. \tag{14}$$

Note that $\phi(\mathbf{k}, \mu)$ is minimum if and only if all the items in $\sup(S_{\mathbf{k},\mu})$ are equal, which implies that $\mu = 1 + \frac{1}{\mu} = g$. So the WM capacity is maximum when μ is exactly the "golden ratio". In this case, it is solved that $a_0 = g^{-1}$ and $b_0 = 0$. Thus, we define a scaling function reflecting the relative lateral inhibition of the dynamical system (1):

$$\phi(\mathbf{k}) = \varphi(\mathbf{k}, \mu = g) = \frac{\inf(\tilde{S}_{\mathbf{k},\mu=g})}{\sup(S_{\mathbf{k},\mu=g})} = g^{\mathbf{k}-2} + g^{-1}. \tag{15}$$

3. Hierarchical Sequential Working Memory Model

Memory model is generally considered to possess a hierarchical structure.[35] In this section, we consider the HSWM model in Fig. 1. More specially, we consider the long term memory as the hierarchical structure of the proposed HSWM. The encoding of working memory can then be considered as the neural bias and the inhibitory weights in HSWM. The proposed hierarchical neural network contains two layers: one PN connected to several CNs, where each PN neuron and its connected CN represents one chunk. It is assumed that there are excitatory connections between the different inhibition connected networks (chunks). These excitatory connections govern the link between PN and CN networks. Once a neuron in PN network becomes a temporal winner, the corresponding CN network can be activated.

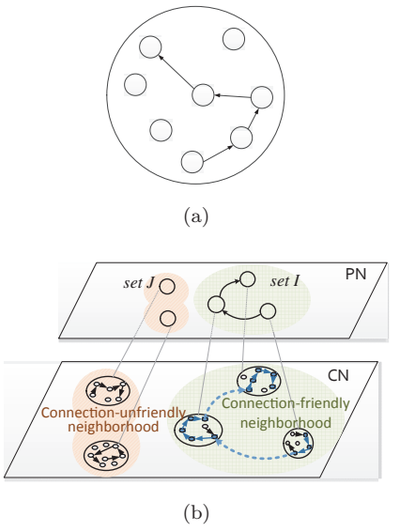


Fig. 1. HSWM model. (a) Path of winner neurons in one chunk; (b) PN and CN networks.

Now we suppose that there are $\mathcal{N}$ chunks $\{\mathcal{C}_1, \ldots, \mathcal{C}_\mathcal{N}\}$ in the HSWM. These $\mathcal{N}$ chunks are controlled by a PN with $\mathcal{N}$ bidirectional inhibition connected neurons. Such a hierarchical structure are described by

$$\begin{aligned} \dot{x}_i &= \tau_{pn} \cdot x_i(\sigma_i - x_i + \Sigma_{j\neq i}\rho_{ij}x_j) + v_i, \\ \dot{x}_{i'}^{\mathcal{C}_i} &= \tau_{\mathcal{C}_{i'}} \cdot x_{i'}^{\mathcal{C}_i}(\sigma_{i'} - x_{i'}^{\mathcal{C}_i} + \Sigma_{j'\neq i'}\rho_{i'j'}x_{j'}^{\mathcal{C}_i}) + v_{i'} \end{aligned} \tag{16}$$

for $i, j = 1, \ldots, \mathcal{N}$ and $\mathcal{C}_i = \{\mathcal{C}_1, \ldots, \mathcal{C}_i, \ldots, \mathcal{C}_\mathcal{N}\}$, $i', j' = 1, \ldots, n_{\mathcal{C}_i}$, where $n_{\mathcal{C}_i}$ is the number of neurons in the corresponding CN $\mathcal{C}_i, \tau_{pn}$ and $\tau_{\mathcal{C}_i}$ are the time constants of PN and CNs, respectively. Usually the time constant $\tau_{\mathcal{C}_i}$ in CN sub networks should be smaller than that in PN network (τ_{pn}). Note that $\tau_{\mathcal{C}_i} < \tau_{pn}$ implies that the dynamic evolving rate in the CN is faster than that in the PN.

Let

$$I_{\text{ch}} = \{\mathcal{C}_{I_1}, \ldots, \mathcal{C}_{I_i}, \ldots, \mathcal{C}_{I_K}\} \tag{17}$$

be set of indices of the chunks which are likely to be the winners among the $\mathcal{N}$ chunks. When the I_ith PN neuron becomes the temporal winner in a time interval $T_0^{I_i} \leq t \leq T_1^{I_i}, \mathcal{C}_{I_i}$ will be the temporal winning chunk during $[T_0^{I_i}, T_1^{I_i})$. At $t = T_0^{I_i}$, the system switches to the CN C_{I_i} which is described by $n_{\mathcal{C}_{I_i}}$ generalized Lotka–Volterra equations. After $t \geq T_1^{I_i} = T_0^{I_{i+1}}$ ($T_1^{I_i}$ and $T_0^{I_{i+1}}$ are the same time point), the next temporal winner in the PN and the corresponding CN is activated.

In each chunk $\mathcal{C}_{I_i}$, the full trace of temporal winners means that all likely to be winner neurons (the number is denoted as $K(\mathcal{C}_{I_i})$) neurons in $\mathcal{C}_{I_i}$ will be winners sequentially. However, practically, an exemplary trace of $\mathcal{C}_{I_i}$ depends on the initial activation state and the time interval during which I_ith PN neuron is the winner. The trace can be denoted as

$$\text{Trace}\{\mathcal{C}_{I_i}\} = \{I_k^{\mathcal{C}_{I_i}} \to I_{k+1}^{\mathcal{C}_{I_i}} \to \cdots \to I_{k'}^{\mathcal{C}_{I_i}}\} \tag{18}$$

for $1 \leq k \leq k'$ and $k' \leq K(\mathcal{C}_{I_i})$, where the initial winning neuron $I_k^{\mathcal{C}_{I_i}}$ depends on the neighborhood to which the initial state belongs. If the initial activation state is

$$\text{Intial}\{\mathcal{C}_{I_i}\} = [0, 0, \ldots, \sigma_{\bar{i}}, \ldots, 0], \tag{19}$$

where $\bar{i} = I_k^{\mathcal{C}_{I_i}}$ is the index of the corresponding neuron in chunk $\mathcal{C}_{I_i}$, Trace$\{C_{I_i}\}$ starts at the neuron $I_k^{\mathcal{C}_{I_i}}$. As long as $T_0^{I_i} \leq t < T_1^{I_i}$, the neural output

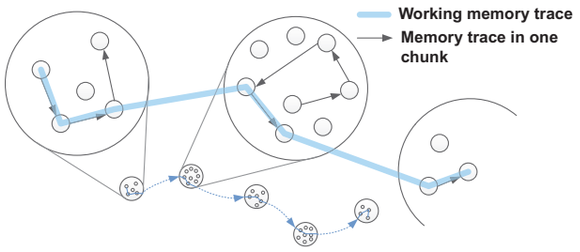


Fig. 2. Memory trace in multiple chunks.

states and the trace order is completely characterized and determined by the equations in (16). At the time $t = T_1^{I_i}$, the trace in $\mathcal{C}_i$ ends at the $I_{k'}^{\mathcal{C}_i}$. When $t \geq T_1^{I_i} = T_0^{I_{i+1}}$, the trace will jump out of $\mathcal{C}_{I_i}$ and go to the next temporal winner chunk $\mathcal{C}_{I_{i+1}}$. Figure 2 illustrates the process of trace of temporal winner neurons in multiple chunks.

4. The Capacity Boundary Analysis

Biological experiments[36,37] validate existence of the upper bound of the inhibition weights between neurons. In Ref. 37, the probability distribution of the ratio of neuron weight and mean weight ($\frac{\rho_{ij}}{|\bar{\rho}|}$) is modeled as a triangle distribution (solid line) shown in Fig. 3, where the weights ρ_{ij} are inhibitory weights within a given chunk (PN or CN). They do not refer to inter-chunk connections, i.e. the connection between a PN network and a CN network. In this paper, we use the triangle distribution to model the inhibition weights ρ_{ij} for all i, j within PN or CN networks. In this case, the mean weight $|\bar{\rho}| = \frac{1}{S} = 2$ where S is the area of the triangle, $0 < |\rho_{ij}| \leq 20$ and

$$\inf(S_{N,\mu}) = \frac{\sigma_{I_N}}{\sigma} + 1 \leq \max(\rho_{ij}) + 1 \leq 21, \tag{20}$$

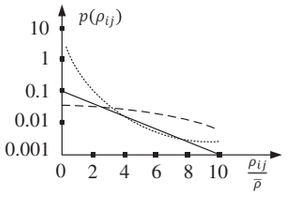


Fig. 3. Distribution of the inhibition weights ρ_{ij} in real neurobiological systems: the dotted lines are the possible distributions from Ref. 37 and the solid line is the distribution we assume in this paper.

where N is the length of the sequence to be encoded. When μ is the golden ratio $g, \sup(S_{N,\mu}) = g$. Let $B = \frac{21}{g} = 12.98$ and then

$$\phi(N) = \left.\frac{\inf(S_{N,\mu})}{\sup(S_{N,\mu})}\right|_{\mu=g} \leq B. \tag{21}$$

This implies that in neurobiological systems, the relative inhibition is less than an upper bound B. This basically underlies the basis of the capacity boundary of working memory.

Now we compare capacity boundary with and without chunking. Form our previous analysis, the scaling function for the non-chunking model is given by Eq. (16). By chunking, a sequence of length N is divided into K chunks. Without loss of generality, we assume that the lengths of the traces in all chunks are equal, i.e.

$$K(\mathcal{C}_{I_1}) = \cdots = K(\mathcal{C}_{I_K}) = N/K. \tag{22}$$

The scaling function reflecting the relative lateral inhibition of the HSWM model becomes

$$\begin{aligned}\phi^K_{\text{ch}}(N) = K(g^{\frac{N}{K}-2} + g^{-1}) \\ + (g^{K-2} + g^{-1}) + c_0,\end{aligned} \tag{23}$$

where $K(g^{\frac{N}{K}-2}+g^{-1})$ and $g^{K-2}+g^{-1}$ are the relative lateral inhibition of CNs and PN, respectively, and c_0 is a constant. Since this reduces to the non-chunking case when $K = 1$,

$$\phi^K_{\text{ch}}(N) = \phi(N) \tag{24}$$

then $c_0 = -2g^{-1}$. Further, we have

$$\phi^K_{\text{ch}}(N) < \phi(N) \quad \text{for some } 1 < K < N \tag{25}$$

and

$$\phi^K_{\text{ch}}(N) > \phi(N) \quad \text{for } K = N. \tag{26}$$

When $K = N$, which corresponds to the case that each item is one chunk, we have $\phi^K_{\text{ch}}(N) > \phi(N)$ implying loss of resources for small chunk sizes. Then the neural circuit can encode a longer trace by chunking for a given value of B. So our proposed model is consistent with the point of view that chunking increases the bound of human WM. For more details, we consider three cases:

Case 1: In one chunk, there is a maximum length N for a given B. In the later simulation, we show that N matches the "magical number 7" as the bound items of human WM when $B = 12.98$.

Case 2: By chunking, there is a maximum length N, which corresponds to an optimal chosen number of chunks K for a given B. We need to solve the problem of finding

$$\max\{N\} \quad \text{subject to } \phi^K_{\text{ch}}(N) \leq B, \ \ K \in \mathbf{N}, \tag{27}$$

where $\mathbf{N}$ is the set of natural numbers. In the later simulation, it is shown that the result matches the "magical number 4" for both the number of chunks and the size of items in each chunk in human WM when $B = 12.98$. Also, the bound of the items of human WM can be increased to 16 items.

Case 3: Assume that B and be greater than 12.98. There is an optimal number of chunks K_0 by minimizing $\phi^K_{\text{ch}}(N)$ with respect to K. By $\frac{d\phi^K_{\text{ch}}(N)}{dK} = 0$, one needs to solve the equation

$$Kg^{\frac{N}{K}} + Kg^K \ln(g) = Ng^{\frac{N}{K}} \ln(g). \tag{28}$$

Suppose K_0 is the optimal solution, then the optimal length of the trace in each chunk is

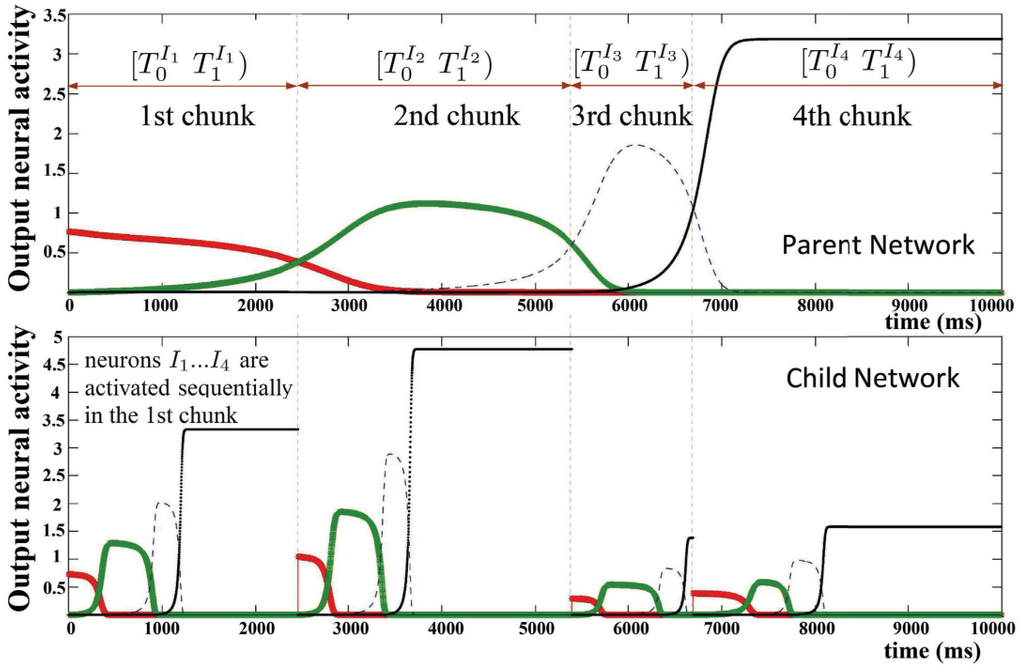
$$K(\mathcal{C}_1) = \cdots = K(\mathcal{C}_K) = \left[\frac{N}{[K_0]}\right]. \tag{29}$$

In the later simulation, it is seen that both the optimal number of chunks $[K_0]$ and the length of trace in each chunk $[\frac{N}{[K_0]}]$ are 4 when $B = 12.98$.

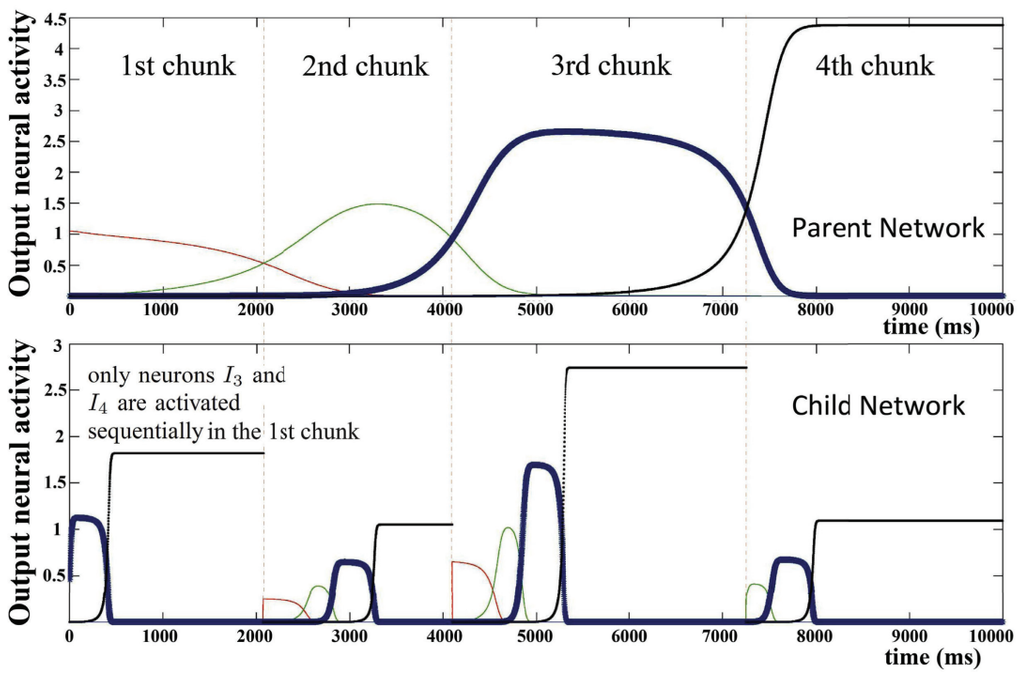
5. Simulation of Memory Trace

In the simulations below, we consider seven chunks ($\mathcal{C}_1, \ldots, \mathcal{C}_7$ with seven neurons in each chunk) in the HSWM model. The neurons in each chunk are divided into two groups: four likely winner neurons (I) and three unlikely winner neurons (J). Suppose that the entry point of each chunk is the neuron with label I_1. In addition, if the chunks have different $\tau_{\mathcal{C}_i}$ and the duration $[T_0^{I_i}, T_1^{I_i})$ lasts sufficiently long, all the likely winner neurons in each chunk will be fully activated. Figure 4(a) shows the full memory trace in the experiments. At beginning, the neuron in the PN that has the maximum output activity becomes the temporal winner neuron and the corresponding CN is activated, which represents the 1st chunk. As time goes on, the (temporal winner neuron) path in the PN is

$$\mathcal{C}_{I_1} \to \mathcal{C}_{I_2} \to \mathcal{C}_{I_3} \to \mathcal{C}_{I_4}. \tag{30}$$



(a)



(b)

Fig. 4. Memory traces of the memory recall: The vertical axis is the output neural activity and the horizontal axis is the sampling index with the sampling time interval $\Delta t = 1$ ms. At a given moment, the neuron which preserves the maximum activity is the temporal winner neuron. (a) Full memory trace: all possible temporal winner neurons are activated sequentially in the CNs of each chunk. (b) An exemplary memory trace: not all the possible temporal winner neurons in CN are activated. This is believed to be a general case in WM recall.

More specifically, in the time interval $[T_0^{I_i}\ T_1^{I_i})$ for $i = 1, \ldots, 4, \mathcal{C}_{I_i}$ will be the temporal winner chunk. And the trace in each the corresponding CN is

$$I_1^{\mathcal{C}_{I_i}} \to I_2^{\mathcal{C}_{I_i}} \to I_3^{\mathcal{C}_{I_i}} \to I_4^{\mathcal{C}_{I_i}}. \tag{31}$$

Note that the time constants (τ) for PN and the CNs are set as 1 and 0.25, respectively. Since, biologically, the probability distribution from $I_{k'}^{\mathcal{C}_{I_i}}$ in $\mathcal{C}_{I_i}$ to $I_k^{\mathcal{C}_{I_{i+1}}}$ in $\mathcal{C}_{I_{i+1}}$ is usually different from person to person and is related to the person's own past experience, the entry point of each chunk is not necessary to be the neuron labeled with I_1. In addition, the subtrace in each chunk in the CNs may jump to another chunk before all the likely winner neurons have been activated. In this case, a memory trace

$$\begin{aligned}\Big\{ & I_3^{\mathcal{C}_{I_1}} \to I_4^{\mathcal{C}_{I_1}} \to I_1^{\mathcal{C}_{I_2}} \to \cdots \to I_4^{\mathcal{C}_{I_2}} \\ & \to I_1^{\mathcal{C}_{I_3}} \to \cdots \to I_4^{\mathcal{C}_{I_3}} \to I_2^{\mathcal{C}_{I_4}} \\ & \to \cdots \to I_4^{\mathcal{C}_{I_4}} \Big\},\end{aligned} \tag{32}$$

which is a subset of the full trace, is shown in the Fig. 4(b).

6. Magical Number 7 and Magical Number 4

Following the analysis for Cases 1–3, our model verifies that sequential WM has a capacity of seven without chunking (Case 1). The biological basis is

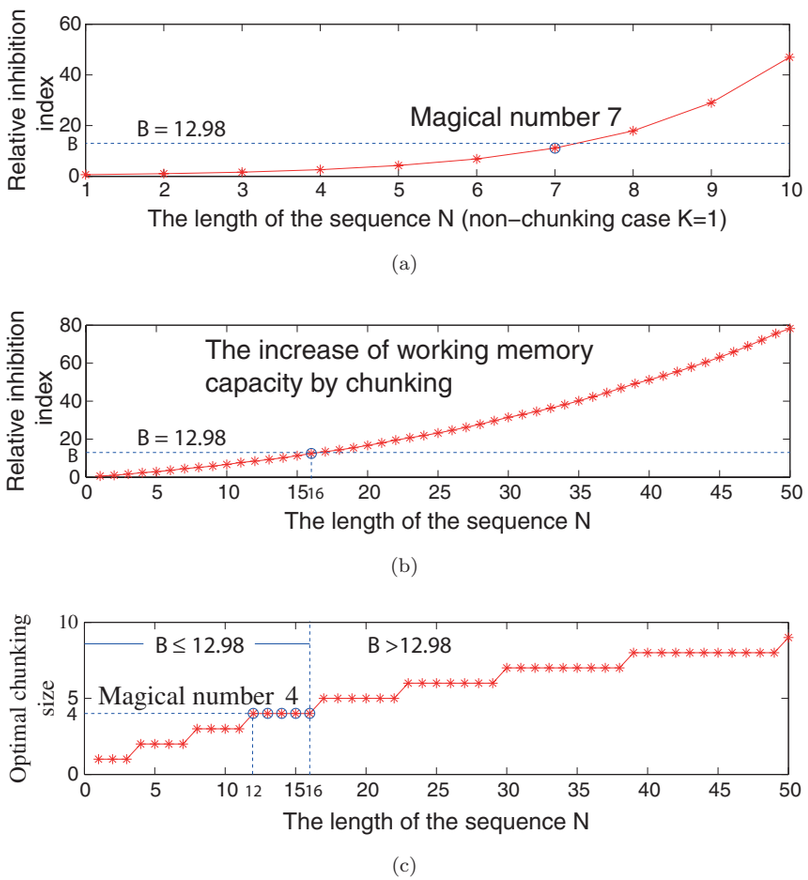


Fig. 5. The "magical number 7" and "magical number 4" in Cases 1–3 in analyzing the capacity boundary of WM. (a) The "magical number 7" without chunking (Case 1). (b) The increase of capacity bound to 16 items by chunking (Case 2). (c) The optimal number of chunks for different length N and the "magical number 4": four chunks and four items in each chunk (Case 3).

that weights in an inhibition connected network are bounded in a very narrow interval (The bound $B = 12.98$ was used in the simulations). The results are shown in the Fig. 5(a). It is interesting to note that the model shows that WM capacity can be increased to 16 items as shown in Fig. 5(b) (Case 2) when the optimal number of chunks is four (see, for example, 12–16 items as shown in Fig. 5(c). In Case 3, we derive that optimal number of chunks $[K_0]$ is expected to be around four with four items in each chunk ($[\frac{N_0}{[K_0]}]$). This is the "magical number 4", reported in recent psychological literature as the capacity of WM. With the analysis above, we have characterized the relationship between the magical numbers 7 and 4, and proposed that this change in capacity is a result of the chunking process.

7. Discussion and Conclusion

Though it is admitted that chunking can increase human WM, theoretical analysis of chunking remains an interesting and open problem. To this end, we have proposed the HSWM chunking model and defined a relative lateral inhibition index for deriving WM capacity. Our paper makes two important contributions to the study of chunking. First, we have shown theoretically that hierarchical chunking can increase the capacity limit of WM. Second, we suggest what chunking changes the capacity limit of human WM from seven items to four chunks, with each chunk consisting of four items (brings the total capacity to 16 items). The additional contribution of this paper is that we discover that HSWM model attains its optimal capacity when the parameter μ in the designed Fibonacci sequence is exactly the "golden ratio". It is to be noted that a possible limitation of the proposed model is that the memory trace generated in (16) cannot intersect itself. Once a neuron becomes a winner in WLC, it cannot become a winner again in the same chunk. For example, when remembering a phone number of "9232-8715", as the digit 2 reoccurs in the same chunk, it needs to be encoded using multiple neurons. Work is in progress to improve the model in order to overcome this issue.

Although the model proposed in this paper has not been experimentally verified, we expect that our proposal will prompt psychologists to design suitable experiments to better understand the capacity limits of WM. Most current studies on WM take into account many factors when considering chunking. These factors, including semantic relations and perceptual features, are dependant on the subject and associations made by the subject through long term memory. As a result, WM is not often analyzed in isolation. Therefore, results often suggest that chunking can increase memory capacity in an unlimited manner through training.[38] For analyzing the role of chunking on the capacity limit of WM it is useful to analyze memory retention shown by amnesiac patients. For instance, one study[14] reports that the upper bound of immediate recall in amnesiac patients is around 15 items.

The HSWM model proposes a model and analyzes the role of chunking in the storage and retrieval of sequential WM. However, such analysis need not be limited to sequential information alone, since, in general, an increase of memory capacity through chunking is a hierarchical process.[39] We therefore believe that our paper lays important groundwork for the theoretical analysis of memory formation and encoding, in both long and short term memory. As such, our model is an important component to be added into memory and cognitive architectures.

Acknowledgments

The authors would like to thank the Science and Engineering Research Council, Agency of Science and Technology Research (A*STAR) for funding the research.

Appendix

Proof of Theorem 1. The proof is based on the investigation of the dissipative saddle point A of $\dot{\mathrm{x}} = \frac{dF(\mathrm{x})}{d\mathrm{x}}$.[40] Let $\mathrm{Re}(\lambda_1) \geq \cdots \geq \mathrm{Re}(\lambda_{r-1}) > 0 > \mathrm{Re}(\lambda_r) \geq \cdots \geq \mathrm{Re}(\lambda_n)$ be the ordered real parts of the eigenvalues of the Hessian matrix $\nabla^2 F(\mathrm{x})$ at A.[41–44] If the saddle value of A defined as

$$\nu(A) = \frac{|\mathrm{Re}(\lambda_r)|}{\mathrm{Re}(\lambda_1)} > 1. \tag{A.1}$$

A is dissipative, which implies that there is a contraction of a deviation after pass its neighborhood. Here we are going to prove that the neurons I_k with coordinates A_{I_k} for $1 \leq k < \mathbf{k}$ are dissipative while $A_{I_\mathbf{k}}$ is a stable equilibrium point of (1).

The encoding process is to design the σ_i and ρ_{ij} for $1 \leq i, j \leq n$. It can be checked that $A_{I_k} = [0, \ldots, \sigma_{I_k}, \ldots, 0]$ is a nontrivial fixed point of (1) and $\nabla^2 F(\mathrm{x})$ is given by

$$\begin{aligned} [\nabla^2 F(\mathrm{x})]_{ii} &= \sigma_i - 2x_i + \sum_{j \neq i} \rho_{ij} x_j \\ [\nabla^2 F(\mathrm{x})]_{ij} &= \rho_{ij} x_i, \quad (i \neq j). \end{aligned} \tag{A.2}$$

In beginning, $\mathrm{x} = A_{I_k}$ and obviously the neuron I_k is the temporary winner. The eigenvalues of $\nabla^2 F(\mathrm{x})$ at A_{I_k} are as follows:

$$\begin{aligned} &\{\sigma_{\tilde{i}} + \rho_{\tilde{i} I_k} \sigma_{I_k}\}_{\tilde{i} \in J} \\ &\{\sigma_{I_i} + \rho_{I_i I_k} \sigma_{I_k}\}_{I_i \in I, i \neq k-1, k, k+1} \\ &- \sigma_{I_k} \\ &\sigma_{I_{k-1}} + \rho_{I_{k-1} I_k} \sigma_{I_k} \\ &\sigma_{I_{k+1}} + \rho_{I_{k+1} I_k} \sigma_{I_k}. \end{aligned} \tag{A.3}$$

We discuss these eigenvalues based on Eqs. (8)–(10) as follows:

Case 1. $\tilde{i} \in J$:

$$\sigma_{\tilde{i}} + \rho_{\tilde{i} I_k} \sigma_{I_k} < -\sigma_{I_k} < 0 \tag{A.4}$$

Case 2. $I_i \in I, i \neq k-1, k, k+1$:

$$\sigma_{I_i} + \rho_{I_i I_k} \sigma_{I_k} < -\sigma_{I_k} < 0 \tag{A.5}$$

Case 3. $I_i = I_{k-1}$:

$$-\sigma_{I_k} < \sigma_{I_{k-1}} + \rho_{I_{k-1} I_k} \sigma_{I_k} < 0 \tag{A.6}$$

Case 4. $I_i = I_{k+1}$:

$$\sigma_{I_{k+1}} + \rho_{I_{k+1} I_k} \sigma_{I_k} > 0. \tag{A.7}$$

If $I_k = I_1$, (9) implies that

$$\begin{aligned} \sigma_{I_1} > \frac{1}{2}\sigma_{I_1} &= \sigma_{I_2} - \left(\frac{\sigma_{I_2}}{\sigma_{I_1}} - \frac{1}{2}\right)\sigma_{I_1} \\ &> \sigma_{I_2} + \rho_{I_2 I_1} \sigma_{I_1}. \end{aligned} \tag{A.8}$$

If $I_k > I_1$, (9) implies that

$$(\rho_{I_{k-1} I_k} + \rho_{I_{k+1} I_k})\sigma_{I_k} > 2\left(\frac{\sigma_{I_{k+1}}}{\sigma_{I_k}} - \frac{1}{2}\right)\sigma_{I_k}. \tag{A.9}$$

For the Fibonacci sequence, we have

$$2\left(\frac{\sigma_{I_{k+1}}}{\sigma_{I_k}} - \frac{1}{2}\right)\sigma_{I_k} = \sigma_{I_{k+1}} + \sigma_{I_{k-1}}. \tag{A.10}$$

Then,

$$|\sigma_{I_k} + \rho_{I_{k-1} I_k} \sigma_{I_k}| > \sigma_{I_{k+1}} + \rho_{I_{k+1} I_k} \sigma_{I_k}. \tag{A.11}$$

Combining above four cases and (A.8)–(A.11), when either $I_k = I_1$ or $I_k > I_1$ with $k \neq \mathbf{k}, I_k$ is dissipative and I_{k+1} will be the next temporal winner since only the eigenvalue $\sigma_{I_{k+1}} + \rho_{I_{k+1} I_k} \sigma_{I_k}$ is positive and its eigenvector direct to $A_{I_{k+1}} = [0, \ldots, \sigma_{I_{k+1}}, \ldots, 0]$. Then, the states will go to the coordinates of the next neuron in the trace

$$\{I_k \to I_{k+1} \to \cdots \to I_{\mathbf{k}}\} \tag{A.12}$$

until it reaches the last one. At $I_{\mathbf{k}}$, all eigenvalues of $\nabla^2 F(\mathrm{x})$ at $A_{I_{\mathbf{k}}} = [0, \ldots, \sigma_{I_{\mathbf{k}}}, \ldots, 0]$ are negative. The coordinate of $I_{\mathbf{k}}$ is a stable equilibrium point of (1). Here we would note that, though the noise is small, it is necessary to avoid the the dynamical system states stopping at an unstable equilibrium.

□

References

1. A. D. Baddeley, Working memory: Looking back and looking forward, *Nat. Rev. Neurosci.* **4** (2003) 829–839.
2. M. Versace and M. Zorzi, The role of dopamine in the maintance of working memory in personal cortex neurons: Input-driven versus internally-driven networks, *Int. J. Neural Syst.* **20** (2010) 249–265.
3. F. Edin, T. Klingberg, P. Johansson, F. McNab, J. Tegner and A. Compte, Mechanism for top-down control of working memory capacity, *Proc. Natl. Acad. Sci.* **106** (2009) 6802–6807.
4. G. A. Miller, The magical number seven, plus or minus two, *Psychol. Rev.* **63** (1956) 81–97.
5. N. Cowan, The magical number 4 in short-term memory: A reconsideration of mental storage capacity, *Behav. Brain Sci.* **24** (2001) 87–114.
6. A. H. Steven and K. S. Todd, Pure short-term memory capacity has implications for understanding individual differences in math skills, *Behav. Brain Sci.* **24** (2001) 124–125.
7. F. Gobet, P. Lane, S. Croker, P. Cheng, G. Jones, I. Oliver and J. Pine, Chunking mechanisms in human learning, *Trends Cogn. Sci.* **5** (2001) 236–243.
8. Z. Chen and N. Cowan, Chunk limits and length limits in immediate recall: A reconciliation, *J. Exp. Psychol. Learn. Mem. Cogn.* **31** (2005) 1235–1249.
9. H. A. Simon, How big is a chunk, *Science* **183** (1974) 482–488.
10. F. Gobet and G. Clarkson, Chunks in expert memory: Evidence for the magical number four or is it two? *Memory* **12** (2004) 732–747.

11. C. Bick and M. I. Rabinovich, Dynamical origin of the effective storage capacity in the brain's working memory, *Phys. Rev. Lett.* **103** (2009) 218101–218104.
12. M. I. Rabinovich, A. Volkovskii, P. Lecanda, R. Huerta, H. D. I. Abarbanel and G. Laurent, Dynamical encoding by networks of competing neuron groups: Winnerless competition, *Phys. Rev. Lett.* **87** (2001) 0681021–0681024.
13. N. Cowan, What are the differences between long-term, short-term, and working memory, *Progr. Brain Res.* **169** (2008) 323–338.
14. A. Baddeley and B. Wilsonb, Prose recall and amnesia: Implications for the structure of working memory, *Neuropsychologia* **40** (2002) 1737–1743.
15. K. Ramanathan, N. Ning, D. Dhanasekar, G. Li, L. Shi and P. Vadakkepat, Presynaptic learning and memory with a persistent firing neuron and a habituating synapse: A model of short term persistent habituation, *Int. J. Neural Syst.* **22** (2012) 12500151–125001520.
16. E. Carter and X. J. Wang, Cannabinoid-mediated disinhibition and working memory: Dynamical interplay of multiple feedback mechanisms in a continuous activator mode of prefrontal cortex, *Cerebral Cortex* **17** (2007) 116–126.
17. B. Poucet and E. Save, Attractors in memory, *Science* **308** (2005) 799–800.
18. C. Johansson, O. Ekeberg and A. Lansner, Clustering of stored memories in an attractor network with local competition, *Int. J. Neural Syst.* **16** (2006) 393–403.
19. G. Li and C. Wen, Convergence of fixed-point iteration for the identification of Hammerstein and Wiener systems, to appear *Int. J. Robust Nonlinear Control*, doi:10.1002/rnc.2837, 2012.
20. G. Li and C. Wen, Convergence of normalized iterative identification of Hammerstein systems, *Syst. Control Lett.* **60** (2011) 929–935.
21. A. Mohemmed, S. Schliebs, S. Matsuda and N. Kasabov, SPAN: Spike pattern association neuron for learning spatio-temporal spike patterns, *Int. J. Neural Syst.* **22** (2012) 1250012.
22. N. R. Luque, J. A. Garrido, J. Ralli, J. J. Laredo and E. Ros, From sensors to spikes: Evolving receptive fields to enhance sensori-motor information in a robot-arm, *Int. J. Neural Syst.* **22** (2012) 1250013.
23. J. L. Rossello, V. Canals, A. Morro and A. Oliver, Hardware implementation of stochastic spiking neural networks, *Int. J. Neural Syst.* **22** (2012) 1250014.
24. T. Yamanishi, J. Q. Liu and H. Nishimura, Modeling fluctuations in default-mode brain network using a spiking neural network, *Int. J. Neural Syst.* **22** (2012) 1250016.
25. W. K. Wong, Z. Wang and B. Zhen, Relationship between applicability of current-based synapses and uniformity of firing patterns, *Int. J. Neural Syst.* **22** (2012) 1250017.
26. A. F. Jahangiri and D. M. Durand, Phase resetting of spiking epileptiform activity by electrical stimulation in the CA3 region of the rat hippocampus, *Int. J. Neural Syst.* **21** (2011) 127–138.
27. A. Vidybida, Testing of information condensation in a model reverberating spiking neural network, *Int. J. Neural Syst.* **21** (2011) 187–198.
28. N. R. Luque, J. A. Garrido, R. R. Carrillo, S. Tolu and E. Ros, Adaptive cerebral spiking model embedded in the control loop: Context switching and robustness against noise, *Int. J. Neural Syst.* **21** (2011) 385–401.
29. S. Ghosh-Dastidar and H. Adeli, Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Integrated Computer-Aided Eng.* **14** (2007) 187–212.
30. S. Ghosh-Dastidar and H. Adeli, Spiking neural networks, *Int. J. Neural Syst.* **19** (2009) 295–308.
31. S. Ghosh-Dastidar and H. Adeli, A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection, *Neural Netw.* **22** (2009) 1419–1431.
32. D. J. Amit and N. Brunel, Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex, *Cerebral Cortex* **7** (1997) 237–252.
33. N. Brunel and X. J. Wang, Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition, *J. Comput. Neurosci.* **11** (2001) 63–85.
34. A. Renart, N. Brunel and X. J. Wang, Mean-field theory of recurrent cortical networks: From irregularly spiking neurons to working memory, *Computational Neuroscience*: *A Comprehensive Approach*, ed. J. Feng (CRC Press, Boca Raton, 2003).
35. A. Tan and H. V. Soon, Concept hierarchy memory model: A neural architecture for conceptual knowledge representation, learning, and commonsense reasoning, *Int. J. Neural Syst.* **07** (1996) 305–319.
36. B. Barbour, N. Brunel, V. Hakim and J. P. Nadal, What can we learn from synaptic weight distributions? *Trends Neurosci.* **30** (2007) 622–629.
37. N. Brunel, V. Hakim, P. Isope, J. P. Nadal and B. Barbour, Optimal information storage and the distribution of synaptic weights: Perceptron versus purkinje cell, *Neuron* **43** (2004) 745–757.
38. K. Ericsson, W. Chase and S. Faloon, Acquisition of a memory skill, *Science* **208** (1980) 1181–1182.
39. L. Feigenson and J. Halberda, Conceptual knowledge increases infants' memory capacity, *Proc. Natl. Acad. Sci.* **105** (2008) 9926–9930.
40. V. Afraimovich, V. Zhigulin and M. Rabinovich, On the origin of reproducible sequential activity in neural circuits, *Chaos* **14** (2004) 1123–1129.

41. G. Li, C. Wen, W. X. Zheng and Y. Chen, Identification of a class of nonlinear autoregressive models with exogenous inputs based on kernel machines, *IEEE Trans. Signal Process.* **59** (2011) 2146–2159.
42. G. Li and C. Wen, Identification of Wiener systems with clipped observations, *IEEE Trans. Signal Process.* **60** (2012) 3845–3852.
43. G. Li, C. Wen, Z. G. Li, A. Zhang, Y. Feng and K. Z. Mao, Model based online learning with kernels, *IEEE Trans. Neural Netw. Learning Syst.* **24** (2013) 356–369.
44. G. Li, C. Wen, G. B. Huang and Y. Chen, Error tolerance based support vector machine for regression, *Neurocomputing* **74** (2011) 771–782.