



Dynamic control of inhibition improves performance of a hippocampal model

S. Polyn, W.B. Levy*

University of Virginia Health System, P.O. Box 800420, Neurosurgery, Charlottesville, VA 22908, USA

Abstract

Adding synaptic modification to the inhibitory interneuron in a minimal computational model of hippocampal region CA3 improves average performance of the simulations. After training on two partially overlapping sequences, simulations are tested on a sequence completion problem that can only be solved by using context dependent information. Simulations with dynamic autonomously scaling (DAS) inhibition are more robust than those without. In the DAS model, scaling factors for inhibition are adjusted gradually over time to compensate for the original model's tendency to move away from a pre-set activity level. This variable inhibition modifies more slowly than the local, associative synaptic modification of the excitatory synapses. As a result, activity fluctuations from one time-step to the next continue to occur, but average activity levels show small variability across training. These results suggest that restricting long term activity fluctuations can be beneficial to recurrent networks that must learn context dependent sequences. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Hippocampus; CA3; Recurrent network; Sequence prediction; Disambiguation

1. Introduction

Fully autonomous recurrent networks are liable to severe activity oscillations. This is particularly true in the more biological models of archi- and neocortex where all the primary recurrent neurons are excitatory. Clearly, there is a role for inhibition as a control mechanism of total network activity. Indeed, anatomical and neurological observations emphasize the importance of this inhibition. For example, a brief anoxic event early in childhood can lead to a small amount of inhibitory neuron death

* Corresponding author.

E-mail addresses: polyn@princeton.edu (S. Polyn), wbl@virginia.edu (W.B. Levy).

Table 1
Summary of network computations

Somato-dendritic excitation/inhibition

$$y_j(t) = \frac{\sum_i w_{ij} c_{ij} z_i(t-1)}{\sum_i w_{ij} c_{ij} z_i(t-1) + K_R \sum_i z_i(t-1) + K_0 + K_I \sum_i x_i(t)}$$

Output

$$z_j(t) = \begin{cases} 1 & \text{if } y_j(t) \geq \theta, \text{ or if } x_j(t) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Typical synaptic modification rule

$$w_{ij}(t+1) = w_{ij}(t) + \varepsilon z_j(t) [z_i(t-1) - w_{ij}(t)]$$

Definitions

y_j	net excitation of neuron j
x_j	external input to neuron j
z_j	output state of neuron j
K_I	feedforward inhibitory scaling factor
K_R	feedback inhibitory scaling factor
K_0	resting conductance
θ	threshold to fire
$c_{ij} \in \{0, 1\}$	variable of connectivity from i to j , kept constant once selected
w_{ij}	excitatory weight, neuron i to j
ε	rate constant of synaptic modification
t	time

accompanied by paroxysmal electrical discharges in the hippocampus. Second, although there are relatively few inhibitory cells in archi- and neocortex (about one in seven in neocortex and even fewer interneurons in the hippocampus), the variety of morphologies and specialized chemistries is quite large. Thus, Nature (natural selection) is willing to spend its genomic currency to control activity levels.

Concurring with these biological observations, it is our experience that controlling activity fluctuations in a model of hippocampal region CA3 is particularly difficult. Indeed, one hippocampal theory of a sequence predicting random recoder [2] could not be implemented in an autonomous network until some knowledge of activity control via appropriate parameterization was achieved [5]. Moreover, this initial version of the model is ultimately unscalable to biological numbers of neurons because uncontrollable oscillations are guaranteed to occur in such a scale up (see [7] especially the appendix).

This scale up problem leads to the introduction of one additional parameter, K_0 (see equation in Table 1). As a result, properly setting the three inhibitory parameters K_0 , K_I and K_R —now allows the model to scale with activity fluctuations kept under control. Thus the model could again function autonomously (our use of autonomous is meant to exclude models with artificial activity control mechanisms such as competitive models).

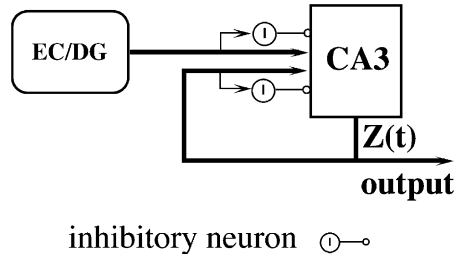


Fig. 1. The hippocampal model. Area CA3 of the hippocampus is modeled by a set of 4096 neurons. A different EC/DG input projects to each neuron of the CA3 system, which, if active, forces that CA3 neuron to fire upon activation of its one and only EC/DG input.

Still, activity control is imperfect, particularly over the course of learning (Fig. 3). Because synaptic weights are modifying, borderline parameterizations can fail due to overly large changes of trial-to-trial average activity. Here we successfully address this problem of large activity changes across training. Moreover, we use the modified theory with its more robust performance to examine an earlier hypothesis: as a random recoder, the CA3 model must be more greatly excited by recurrent connections than by the feedforward connections [2,3,6].

To control the large activity shifts that can accompany learning, we hypothesize inhibitory synaptic modification that compensates for a changing recurrent excitation. We call this process dynamic autonomously scaling (DAS) inhibition. The inhibitory mechanism implemented here uses a local, error correcting mechanism implemented on a time-scale slower than the moment-to-moment updating of excitatory neuronal states [1]. Thus due to the time scale selection, DAS restricts—but does not abolish—activity fluctuations. Interestingly, the addition of this complexity to the model actually allows a reduction in the free variables in the model.

2. Methods

The network, a 4096 (except as noted otherwise) neuron model of region CA3 of the hippocampus, is a sparsely connected (10%), recurrent excitatory system (see Table 1). In addition to the recurrent connections, each McCulloch-Pitts neuron receives a single external input that represents its entorhinal cortex and its dentate gyrus (EC/DG) inputs. This single input is hypothesized to be very powerful if activated (Fig. 1). There are also two inhibitory interneurons that control activity fluctuations in a non-precise way (see Fig. 1 and Table 1).

The model uses a post-synaptic associative modification rule (Table 1) that is disabled during testing. The activity of the network (the number of neurons active on a given time-step) is analyzed in terms of the fraction of external inputs $I(t)$ and the fraction of neurons fired by recurrent activity $R(t)$ on that time-step [6]. Without the time variable, the term I refers to the average $I(t)$ across all time-steps for that trial. Activity fluctuations are controlled by two inhibitory interneurons and

Table 2
Dynamic autonomously scaling inhibition

$$m_a = \frac{\sum_{t=1}^{12} \sum_{i=1}^{4096} z_i(t)}{12}$$

$$K_R(\tau) = K_R(\tau - 1) + 0.8 * m_a * (m_a - m_d)$$

m_a is the average activity for the network over a single trial

4096 is the number of neurons in the simulation

12 is the number of time-steps in a single training trial

$z_i(t)$ is the firing value of a given neuron i at time step t . $z_i(t)$ is binary

$K_R(\tau)$ is the coefficient of feedback inhibition at training trial τ

0.8 is a rate constant

m_d is the desired average activity level

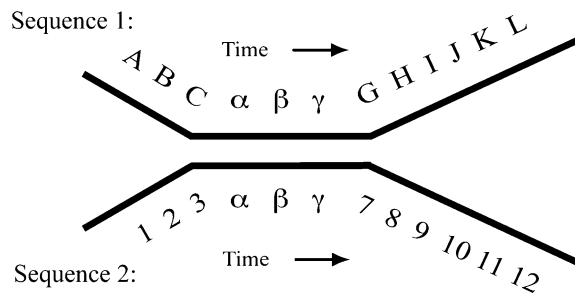


Fig. 2. A schematic of the input sequences. Two input sequences are presented to the network. Each sequence is composed of 12 patterns. The EC/DG input patterns are broken up into three orthogonal subsequences—the initial subsequence, the shared subsequence and the tail subsequence. A pattern consists of a set of neurons being turned on for one time-step. The patterns within a subsequence have a slow shift rate, that is, each pattern has 85% of its neurons in common with the patterns preceding and following it.

a resting conductance (K_R , K_I , and K_0 , respectively). Overall network activity is non-competitive, and only roughly controlled by the inhibitory scaling constants (Table 2).

The inhibitory interneuron in the model has access to the total activity of the principal cells. All three of the inhibitory scaling factors are modified by the DAS equations. K_I and K_0 are changed in the same proportion that K_R is changed. That is, the initial proportion between K_R and K_I , and K_R and K_0 is preserved.

Each training trial consists of a random input pattern followed by one of two sequences, each 12 patterns long. The initial randomization of the network state, a binary vector of the desired activity level, varies from trial to trial. Each individual input pattern of the sequenced input pattern corresponds to a set of neurons in the network that are activated for a single time-step by an external connection. As Fig. 2 shows, each of the two input sequences contains three orthogonal subsequences. Notably, the two sequences share a common, central subsequence of three patterns (α , β , and γ of Fig. 2). There are 350 presentations of each sequence, with the presentation of the two sequences interspersed.

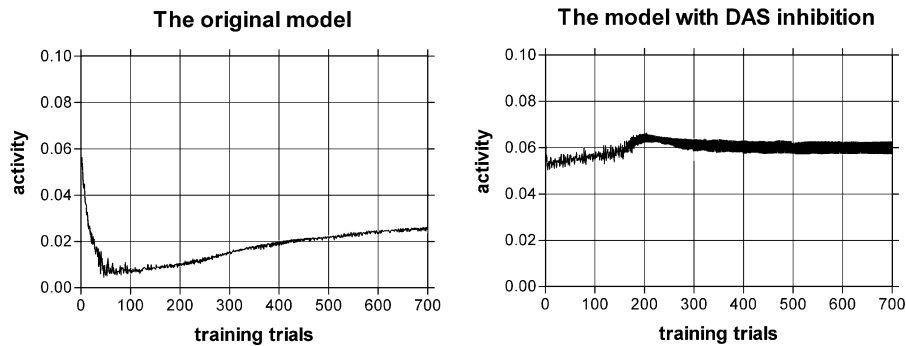


Fig. 3. Activity across training. Each point represents the average network activity over a single trial. Each trial consists of 12 time-steps. (A) The activity falls to about 15% of its original value over the first 50 trials. This is due to changes in the weights of the network due to learning. (B) The model with DAS inhibition is able to compensate for changes in the weights of the network. The inhibitory scaling constants are modified by a simple error correcting equation to keep activity relatively constant.

Testing consists of sequence completion in response to a partial input. A cosine comparison (the normalized dot product of two vectors) is used to decode each network state that evolves in response to the test input. The cosines compare each individual state produced by the two input sequences at the end of training with states evolved during testing. By definition, the largest value at each time-step is the decoded network state at that moment. The networks are tested on the disambiguation task. From each starting point (1 or A of Fig. 2), the network must predict the appropriate path to the end of the learned sequence (12 or L, respectively). In this problem, the test input is only activated for a single time-step.

For each pair of specifications—{total activity and external activity}, there are 10 simulations with each simulation differing by its random connectivity. Each of these 10 trained networks is tested 10 times.

3. Results and discussion

Using the original model lacking DAS inhibition, the activity levels shift 2–3 fold across the first 100 training trials. For example, one simulation begins at 6% activity and drops to 2% (Fig. 3a) while another starts at 7% and rises to 13%. On the other hand, with the DAS modification there is only a small dip or rise in average activity (Fig. 3b).

In [6], it was determined that there is an optimal level of I , external activity, to perform the disambiguation problem ($0.1 < I < 0.2$). In the present study we were able to determine that this optimal ratio of feedforward and feedback activity holds not only when the network scales up, but when activity is lowered to more biologically plausible levels as well. Fig. 4 shows how this modification helps performance. At 6% activity and 4096 neurons in the network, the original model was unable to perform

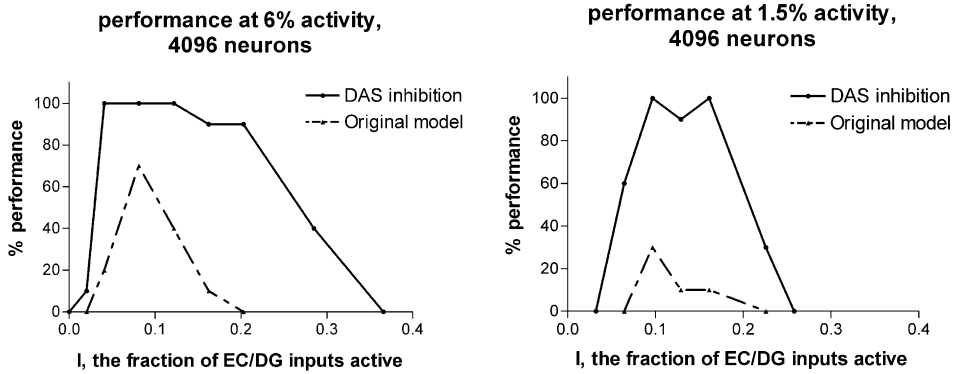


Fig. 4. DAS inhibition increases robustness of performance as a function of external activity. At 6% activity and 4096 neurons, the original model never reaches 100% performance. Fractional activity is the proportion of firing neurons that are activated by the EC/DG input. Performance measures the number of simulations out of 10 that are able to successfully solve the disambiguation problem. (B) Activity can be lowered to 1.5% with DAS inhibition. Approximately 61 neurons fire on each time-step out of 4096. The original model is severely impaired in its performance at this activity level.

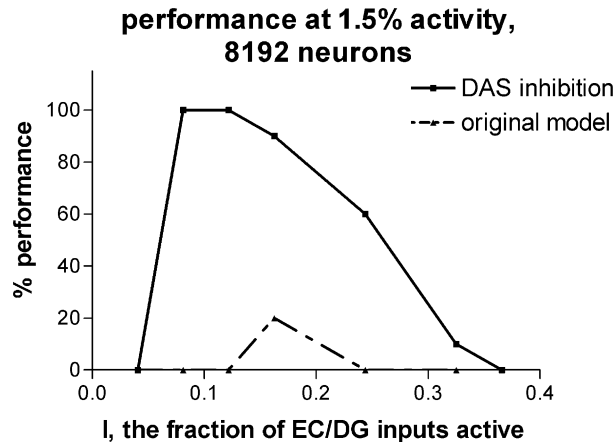


Fig. 5. DAS inhibition allows network size to scale up to previously unreachable values. At 1.5% total activity and 8192 neurons there is still a region of perfect performance. Further scale up was prevented by computational constraints.

perfectly at any level of I . With the DAS modification, there is a range of I in which the model exhibits perfect or near-perfect performance. Even with the average activity lowered to 1.5%, there is an area of perfect performance. Scaling the model up to 8192 neurons (Fig. 5) the optimal level of I is still around 0.1.

The DAS inhibition model changes the failure mode for the simulations that are unsuccessful at disambiguation. Instead of activity spontaneously dropping to zero,

now failing simulations maintain activity but do not learn to distinguish the two sequences.

In the original hippocampal model, the three inhibitory scaling factors (K_R , K_I and K_0) were free variables. Presently, we set the activity level, and use the equations from [7] with *gradient set to zero* to calculate the corresponding K values. In this incarnation of the model, activity level is the only free variable.

4. Increasing the robustness of the model

The addition of DAS inhibition to the hippocampal model makes it remarkably more robust. This robustness is exhibited in two ways. First, it allows us to use larger numbers of neurons in our simulations, without worrying about destructive activity fluctuations. This allows us to approach more biologically relevant network sizes (8192 neurons). Second, it allows us to lower activity to more realistic levels (less than 2%). What was before an impossibly low activity is now easily achievable by the model.

It is important to note that the DAS inhibition model bounds activity in a non-destructive fashion. Here large scale fluctuations are minimized while, in contrast to competitive activity control, time-step to time-step fluctuations are allowed to occur. Thus the more biological model outperforms the less biological model with its perfect activity control. (For the destructive effects of competitive activity control, see [4].)

Acknowledgements

This work was supported by NIH MH48161 to WBL and MH57358 to N. Goddard (with subcontract #163194-54321 to WBL), and by the Department of Neurosurgery.

References

- [1] R.A. Chitwood, A. Hubbard, D.B. Jaffe, Passive electrotonic properties of rat hippocampal CA3 interneurons, *J. Physiol.* 515 (3) (1999) 743–756.
- [2] W.B. Levy, A computational approach to hippocampal function, *The Psychology of Learning and Motivation*, Vol. 23, Academic Press, San Diego, 1989, pp. 243–305.
- [3] W.B. Levy, A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks, *Hippocampus* 6 (1996) 579–590.
- [4] W.B. Levy, X.B. Wu, Some randomness benefits a model of hippocampal function, in: H. Liljenstrom, P. Arhem, C. Blomberg (Eds.), *Disorder versus Order in Brain Function*, World Scientific Publishing, Singapore, 2000, pp. 221–237.
- [5] A.A. Minai, W.B. Levy, Setting the activity level in sparse random networks, *Neural Comput.* 6 (1994) 85–99.
- [6] S. Polyn, X.B. Wu, W.B. Levy, Entorhinal/dentate excitation of CA3: a critical variable in hippocampal models, *Neurocomputing* 32–33 (2000) 493–499.
- [7] A.C. Smith, X.B. Wu, W.B. Levy, Controlling activity fluctuations in large, sparsely connected random networks, *Network: Comput. Neural System* 11 (2000) 63–81.