

Defining Time in a Minimal Hippocampal CA3 Model by Matching Time-span of Associative Synaptic Modification and Input Pattern Duration

Kurt E. Mitman¹, Patryk A. Laurent¹, William B Levy²

Department of Neurosurgery², College of Arts and Sciences¹, University of Virginia

P.O. Box 800420, Charlottesville, VA 22908

Abstract – This paper quantifies the time shifting of neuronal codes in a sparse, randomly connected neural network model of hippocampal region CA3. As this network is trained to learn a sequence, the neurons that encode portions of this sequence characteristically fire earlier and earlier over the course of training. Here we systematically investigate the effects of the N-methyl-D-aspartate(NMDA)-governed time-span of synaptic associativity on this shifting process and how this time-span interacts with the duration of each successive external input. The results show that there is an interaction between this synaptic time-span and externally applied pattern duration such that the early shifting effect approaches a maximum asymptotically and that this maximum is very nearly produced when the e-fold decay time-span of synaptic associativity is matched to the duration of individual input patterns. The performance of this model as a sequence prediction device varies with the time-span selected. If too long a time-span is used, overly strong attractors evolve and destroy the sequence prediction ability of the network. Local context cell firing – the learned repetitive firing of neurons that code for a specific subsequence – also varies in duration with these two parameters. Importantly, if the associative time-span is matched to the longevity of each individual external pattern and if time-shifting and local context length are normalized by this same external pattern duration, then time-shifting and local context length are constant across simulations with different parameters. This constancy supports the idea that real time can be mapped into a network of McCulloch-Pitts neurons that lack a time scale for excitation and resetting.

I. INTRODUCTION

The sequence learning, recoding theory of hippocampal function [1] depends heavily on a temporally asymmetric rule governing associative synaptic modification. This temporal asymmetry plays a critical role for a special kind of sequence completion: it allows learning that produces timely predictions [2]. That is, it enables the network to meet an important requirement – use an early part of a sequence to generate the rest of the sequence before the rest of the sequence occurs. The asymmetric associative temporal characteristic of synaptic modification arises from the on- and off-rate properties of the NMDA receptor [3].

The recoding process itself has been predicted as time-shifting to earlier firings [1,4] resulting in a somewhat loosely overlapping code (see “rough-counting” in [1]). In this way a sequence of neural firing resembles an incrementing, shifting binary counter when the neurons are visualized via re-ordering by the temporal order in which they fire. In terms of biology, the earlier or backwards shifting of firing in the

model is similar to that recently demonstrated in vivo [5] where it was complemented by a negative skew in the activity of a large portion of the neurons [6]. The combination of earlier-shifting of firing with the formation and lengthening of place-cell type firing results in a neural code for the present input which becomes increasingly similar to the future patterns. This similarity is what allows timely predictions [1].

When the model learns a nonspatial task, cell firing resembling place cells occurs, and these neurons are called local context neurons because they identify a particular subsequence of a longer sequence. The average duration of such place-cell type firings is denoted $E[L]$, the average local context length. When training the model to learn cognitive tasks, an intermediate to high value of $E[L]$ correlates with good performance (e.g., [7,8]). However, if $E[L]$ becomes excessively large, the network state can be drawn into a noisy attractor, which could prevent the appropriate sequence recall. This motivates us to study $E[L]$ as a function of parameters such as the time constant of synaptic associativity. As an expected result, we find that $E[L]$ increases with increasing values of the time constant of associativity.

A central finding of this paper is a matching between the time-span of associative modification and the duration of the inputs (stutter length) to the network. We show that this matching produces a compromise between the predictive component of the neuronal codes developed during learning and performance. *Specifically increasing the LTP time-span beyond this compromise does in fact marginally increase the predictive component of the neuronal codes, but it also degrades performance and the robustness of the codes.* That is, too large of a time-span can cause the formation of performance-destroying stable attractors. The other important observation here concerns the mapping of time between computational cycles of the simulations and real time.

By ratioing against input pattern longevity (stutter length), the neuronal codes developed by training are constant across parametric changes of the NMDA receptor off-rate time constant that maps the model into real time. Thus, the longevity of neuronal firings and the time shifts measured are also mapped into real time.

II. METHODS

A. The Model

The model simulates the CA3 region of the hippocampus, an area sometimes thought to be the center of associative memory due to the presence of recurrent connectivity [9,10]. Because we implement an asymmetric synaptic modification rule [11], the CA3 region is also able to make associations across time. Therefore, context-dependent sequence learning takes place in this model (see [12], for a review).

The model reflects the biology of the hippocampus, yet at the same time aiming for easy interpretation via minimality. The network is comprised of McCulloch-Pitts binary neurons. The recurrent connectivity between neurons is random and sparse. The connections to the CA3 region from the entorhinal cortex and dentate gyrus are combined and represented as a single input vector. In contrast to the visual model, here inhibition is implemented using a *k-winners-take-all* competitive rule. The following equations illustrate the dynamics of the model:

For neuronal excitation,

$$y_j(t) = \sum_{i=1}^n c_{ij} \times W_{ij} \times Z_i(t-1);$$

for output,

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

where $y_j(t)$ is the net excitation for the j^{th} neuron at time t , for $j \in \{1, \dots, n\}$ neurons, $Z_j(t)$ is the output, $W_{ij}(t)$ is the weight of the connection from neuron i to j ; $c_{ij} \in \{0, 1\}$ is the connection indicator, $x_i \in \{0, 1\}$ indicates which neuron are activated by the entorhinal cortex/dentate gyrus (externals), and θ is the threshold, determined each time step such that only the 7.5% of neurons with the highest activity have an output of 1.

Synaptic weights are modified in accordance with a temporally asymmetric Hebbian rule [11,13,14] and have a time-spanning associative capability that models a saturation-decay model of the NMDA receptor. The equations are as follows:

For time-span associativity a saturate and decay variable $\bar{Z}_i(t)$ is used,

$$\bar{Z}_i(t) = \begin{cases} 1 & \text{if } Z_i(t) = 1, \text{ and} \\ \alpha \times \bar{Z}_i(t-1) & \text{otherwise} \end{cases}$$

for synaptic modification,

$$W_{ij}(t+1) = W_{ij}(t) + \mu Z_j(t) (\bar{Z}_i(t-1) - W_{ij}(t)).$$

The term μ in the weight modification equation is the learning rate constant for weight changes and depends on stutter length, $\mu = [(1.05)^{1/\text{stutter}} - 1]$. The term α in the presynaptic updating of $\bar{Z}_i(t)$ is the decay time constant of the NMDA receptor. We assume e -fold decay in the NMDA receptor after 100 ms. We can represent this decay in discrete time as the number of time steps it takes for \bar{Z}_i to reach $1/e$ if no further excitation arrives on input line i . Thus, the length of real time represented by each time step (in ms) = $-100 / [\log(\alpha) \times e]$.

All simulations used 8% connectivity (e.g. $P(c_{ij} = 1) = .08$), where self connections are not allowed and $n=4096$. This value of connectivity is inferred from the projections of the roughly 83,000 neurons in the septal one-third of CA3 of the rat hippocampus. Random synaptic transmission failures occur at each synapse at a rate of 20%. The results presented, however, are not significantly different from those where synaptic transmission failures were absent from the network model.

B. Inputs

The network was trained on sequences of ten orthogonal (non-overlapping) patterns. Each pattern was presented for multiple time steps. The number of time steps each pattern is activated is called “stutter length.” An input pattern on one time step activated 64 neurons randomly chosen from sets of 96 neurons. That is, on each time step, a random 67% of the neurons of the appropriate external input pattern fire (see Fig. 1, top graph).

There are two types of sequences which are simulated: training trials, in which the entire externally driven sequence is presented and synaptic modification is allowed to occur and testing trials, in which part or all of the sequence is presented and no synaptic modification is allowed to occur.

C. Determining Shift

Shift is a measure of the change in temporal position of neural firing as a function of training. The metric used here compares neuronal firing after an “early” training trial to the final training trial. The “early” training trial is determined to be that when the network has achieved a stable initial firing distribution across time-steps even though firing is still mostly random at the beginning of training due to the random connectivity (See Fig. 1). We find that after 5 training trials most neuronal initial firing times can be reliably predicted across the sequence. We ignore, however, neurons from the final two patterns of the sequence because of a nonrepresentative end of sequence effect.

When α is non-zero, shift always tends to be in the earlier direction for the neurons in a network. We do find that some neurons which originally fired at the beginning of sequences appear to shift later (towards the end of the

sequence) by large numbers of time steps. This “later shifting”, however, is unrelated to the earlier-shifting phenomenon being studied. We hypothesized these neurons, which can later-shift by as much as 75% of the sequence, are inefficient coding neurons from the early patterns which through the process of synaptic modification are depotentiated, cease firing, and then are recruited later to code for patterns in the middle or at the end of the sequence. Also, in the case of neurons which have inputs from externals as well as recurrences driven by those same externals marginal forward shifting is possible, representing a different mechanism. A more detailed analysis of these other mechanisms is still necessary.

D. Measuring Performance

Successful learning is defined as a simulation being able to complete 80% of the training sequence during recall testing (only the first input pattern is externally activated) and doing so without skipping more than two patterns and without reaching an inappropriate stable state before the end of the sequence (see Fig. 2).

To determine which external pattern is being represented at each time step of recall, the neuronal firing is compared to the previously recorded neuronal firing on every time step of training (where all patterns were externally activated). The time step of training which is most similar to that of the network state at time t of recall (as determined by an inner product of the firing) is the input represented by the network during recall at time t .

E. Measuring $E(L)$

Local context firing length measures the contiguous period of time over which neurons fire. Published findings from our laboratory show that an intermediate to high average value of local context length, $E(L)$, is a good predictor of learning (e.g., [7,8]). We measure local context length using externally driven sequences and then count the period of time from a neuron’s first time-step of firing to its last time-step of contiguous firing. Temporally contiguous firing is defined as two firings with less than three non-firing time-steps between the firings (e.g. on-off-off-on are temporally contiguous, on-off-off-off-on are not).

Because failure prone synaptic transmission and noisy inputs are being used here, some neurons flicker on-off-on while still being a local context neuron. The entire length was used to value such a neuron (e.g., on-off-on implies a length of 3). Some other neurons are randomly activated by the random processes just mentioned and are not strongly associated with any particular subsequence, or time-step, of the training sequence. These neurons were characterized by wide gaps between firing times and lack of temporally adjacent firings. Such neurons are categorized as unused.

For the first two or three training trials neural firing is almost entirely random. However, by trial five it is possible to reliably identify a particular time step with the firing of a particular neuron. With additional training trials the firing of the neurons begins to shift earlier in the sequence. Earlier shifting is clearly demonstrated by the externally activated neurons (see Fig. 1). The average shift of the externally activated neurons at the end of training is around fourteen time steps or 154 ms or just over one and a half patterns.

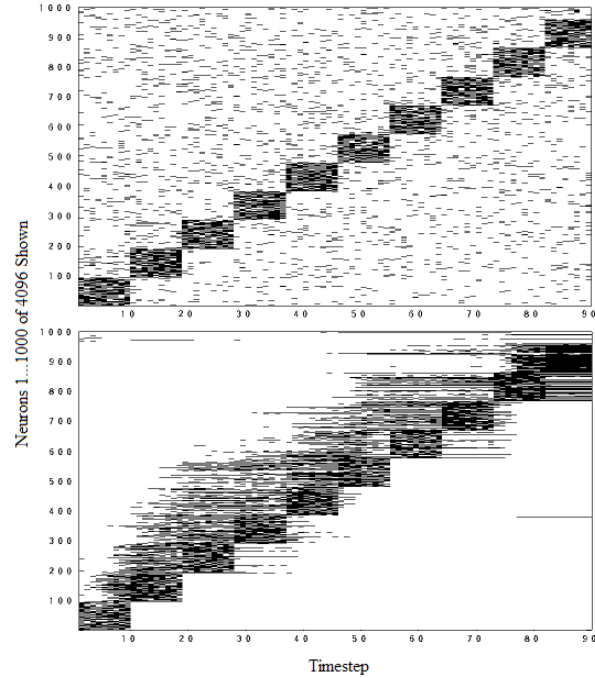


Fig. 1. **Training produces time shifting to earlier positions of externally activated neurons by extending context length.** Firing of the first 1000 neurons of the 4096 total neurons in a network during the fifth (upper figure) and 250th (lower figure) training trials. The external inputs to the network include neurons 1 through 960. Neurons 961 through 1000 are recurrently activated if they fire. The inputs are a sequence of orthogonal patterns, each on for nine time-steps (stutter=9). The rectangular blocks along the diagonal in the upper figure are the externally activated neurons. These rectangles are not solid because of input noise that randomly deselected external activation.

Typically, however, the time-shifting of the externally activated neurons tends to be less pronounced than that of purely, recurrently-activated neurons. Because an externally activated neuron is turned regardless of other activity, there is strong depotentiation of the recurrent inputs to these neurons, synaptic weakening beyond a reclaimable weight value. Thus, such a neuron tends toward a functional disconnection from most other neurons, thereby decreasing the likelihood that it can be activated earlier.

III. RESULTS

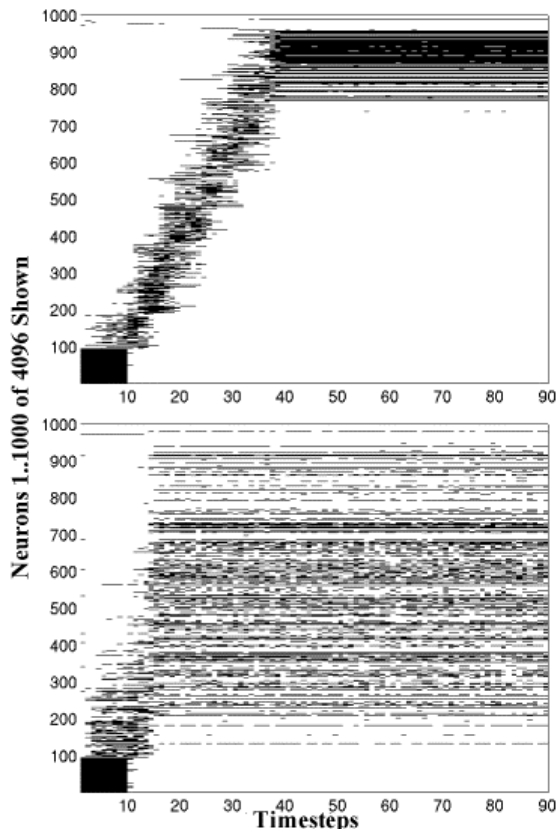


Fig. 2. Example of a simulation that satisfies (top) and fails (bottom) to meet the 80% learning criterion on recall. Both graphs are testing trials after 250 training trials. During the testing trial the first pattern is presented (dark rectangle) and then the network is allowed the free run. The top graph has a matched value of α to the stutter (note, on recall the sequence is compressed by a factor ~ 2). The bottom graph has a value of α above the matching value and falls into a steady state attractor.

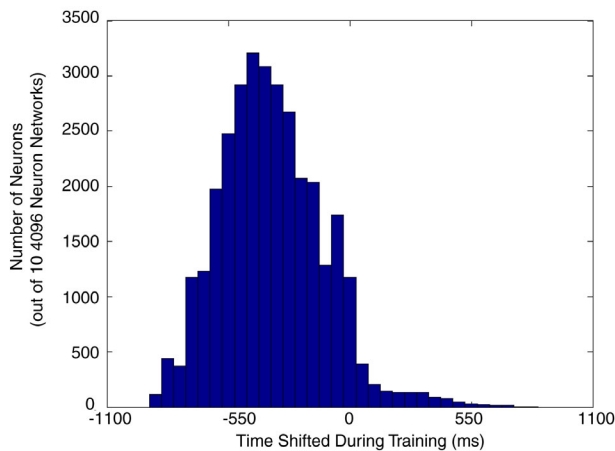


Fig. 3. Time shift histogram of the recurrent neurons in 10 networks with externally activated neurons excluded. Shift for each neuron is defined as the change in onset time of its initial firing after trial 5 vs. trial 250 (see Methods). Temporal shifting in the network during learning is predominantly in the earlier direction. The median total shift is approximately -34 time steps. However, this figure does not take context length, $E(L)$, into account, and so does not fully reflect the skewing of place fields. The stutter was 9 and α was matched to this stutter length.

The earlier shifting seems to exhibit asymptotic behavior, that is by trial 150 neurons are reliably firing to a particular subsequence of the training sequence and shift only slightly as training proceeds. The average number of time steps shifted at the end of training is quite sensitive to the value of α , particularly when $-\log(\alpha) > 1/\text{stutter}$ (See Fig. 3). While there is no local minimum at e-fold decay ($\alpha^{\text{stutter}} = e^{-1} \approx 0.3679$), beyond e-fold decay rates, earlier-shifting approaches an asymptotic value (See Fig. 4). The mean amount of earlier-shifting appropriately reaches the maximum backward shift by setting the time-span of associativity to the e-fold decay value suggested by biology. The results show that the e-fold decay rate yields the maximum earlier-shifting with the minimal amount of code and learning degradation (only 15% of networks fail to complete the sequence on recall). Increasing α even slightly beyond the e-fold decay rate results in network failure in more than 50% of networks (See Fig. 4). Therefore, matching the time-span of associativity to the input pattern duration (in the model, matching the parameter alpha to the parameter stutter) is arguably a useful compromise for the model and, as a hypothesis, explains the similarity between the time of a theta cycle and the NMDA-R off-rate time constant in rats.

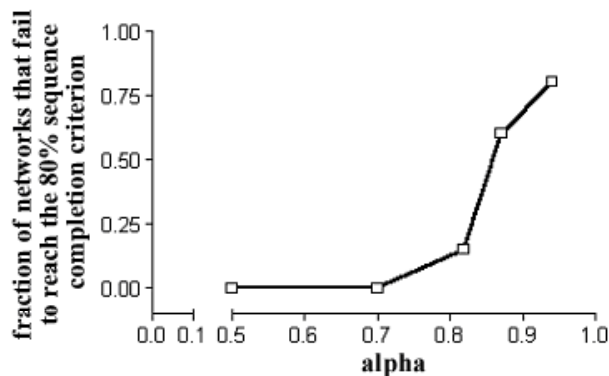
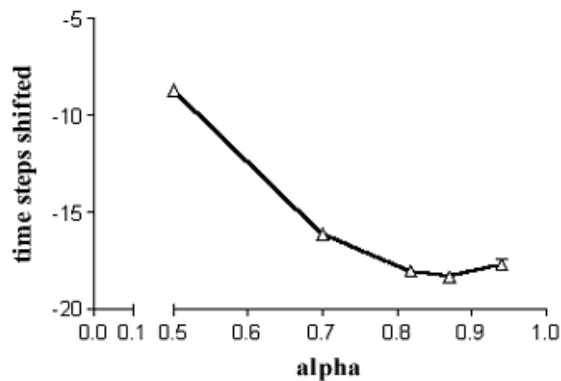


Fig. 4. The effects of α on performance and shift at stutter-length five. When α is matched to stutter-length ($= 0.82$), the time shift (upper figure) has essentially achieved its maximum backward value. However, at this value an overly strong attractor is beginning to develop to the detriment of performance (lower figure), at least for the competitive networks used here. Each point represents the average of 20 randomly connected networks.

At the beginning of training, the randomness in connectivity and random neural firing imply that, besides the externally activated neurons, almost no place-cell type firing (local context firing) is present in the network. However, after five training trials, such local context firing begins to emerge. These regular firings are typically on the order of one or two time steps in duration. With repeated training trials the firing length for context neurons increases, and non-context neurons begin to stop firing altogether. We find that $E(L)$ at the end of training is sensitive to α , increasing almost linearly with increasing α . This result is closely related to the earlier-shifting of neurons. That is, neurons not only begin their context firing earlier in each trial as training progresses, but they maintain part of their previous firing times. As a result, overall local context firing times lengthen for such neurons. At the matching e-fold decay rates, $E(L)$ is typically 1.9 times the input pattern duration (see Fig. 5).

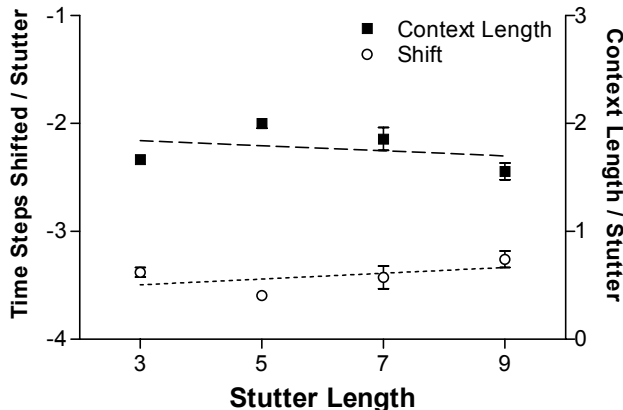


Fig. 5. **Normalized time shifting and normalized context length are constants.** Mean time shifting divided by stutter length and median context length divided by stutter length are plotted for four different values of stutter with each stutter length matched by the e-fold, decay constant α . Because a constant value is produced, time may be physically and consistently represented in the model by the ratioing and matching α to stutter. Each point is the average of ten random networks. The slopes of the two regression lines is approximately 0.01 and do not represent a significant deviation from zero in the slope.

The meaningfulness of matching stutter length to the associative time span is emphasized by normalized measurements of average time shifting and of average context length and is illustrated by the mean zero slope lines of Fig. 5. That is, if we divide training-induced average context length by stutter length when training at the matched time span, the model becomes parameter independent. Thus the matching of stutter length and time span becomes an explicit definition of time in the model because the associative time span is measurable and available from experimental work. Moreover, the normalized shift value and the normalized context length, $E(L)$, are a physiological prediction if we assume one complete pattern cycle is one complete theta cycle. And there is another prediction: slow animals – like

humans – have a slower NMDA-R off-rate time constant in CA3 than fast, small animals like rats and mice.

IV. DISCUSSION

Most of these paragraphs address the origin of the earlier-shifting phenomenon across training. The last paragraph points out a general hypothesis about the evolved function and parameterization of the hippocampus.

Because of the direct externally activated excitation and because of indirect external excitation through recurrent neurons, any one recurrent neuron will be variously biased to fire as a function of time within a training trial (at the outset of training). Because the external activation is stuttered, the bias of any one, particular recurrent neuron is also stuttered. That is, an approximately constant bias covers a more-or-less contiguous time interval equal to the stutter-length for such a neuron. Now hypothesize random firing of recurrent neurons reflecting such biases, then we claim that the temporally asymmetric synaptic modification rule produces a tendency for such neurons to fire earlier over the course of training for two reasons:

First, because of the temporal asymmetry of the synaptic modification rule, only those inputs turning on somewhat before a postsynaptic neuron fires, or would tend to fire due to a strong bias, will on average potentiate. Contrariwise, those inputs tending to turn on afterward or not turn on at all will, on average, have their synapses weakened. Thus, the inputs that get strengthened are the ones that turn on earlier. But again recall, these inputs too are biased to turn on for several steps just as the relevant postsynaptic neuron is biased toward firing for several time-steps. So as these inputs strengthen, the postsynaptic neuron will begin to fire earlier. These ideas, however, need to be generalized.

Consider now two sets of equal biases that excite the same postsynaptic neuron but are at different noncontiguous places in the sequence. Just by chance one bias will grow stronger than the other because of synaptic modification and the random aspects of the inputs. Then this stronger bias will tend to wipe out the weaker bias by virtue of the synaptic modification rule weakening the synaptic inputs through which the weaker bias is expressed.

Second, if any one particular neuron is biased strongly enough so it begins to fire with regularity in response to a particular subsequence, this neuron will contribute to the bias of other neurons just as if it were an externally activated neuron. The fact that this particular neuron is shifting earlier, in accordance with the reasons just above, results in a *cascade* of earlier shifting. That is, as a set of temporally associated neurons shift earlier so will the neurons they activate. Thus, it makes sense to refer to an earlier-shift or backward cascade that can propagate through assemblies of coding neurons over successive trials.

The earlier shifting of neurons can be controlled by the value of α . We take α to be the rate constant for the unbinding of glutamate from the NMDA receptor. We

propose that the role of hippocampal recoding is a compromise between maximizing the predictive value of the codes developed while preserving the ability of the CA3 to sequence. However, as we increase α we find that there is a point at which overly strong associations create noisy stable attractors, destroying the sequencing capabilities of the sparsely connected recurrent network.

Neurophysiological findings are consistent with an e-fold decay rate in which the presynaptic neuron's contribution to LTP decreases. Presumably the e-fold decay rate arises from the unbinding of a glutamate saturated NMDA-R system. As such, we can write the unbinding as proportional to $e^{-t/\text{constant}}$, for some constant that we set to one. Suppose at $t = 0$, the saturated system has a value of one. Then on the next timestep, Δt , its value is α ; that is, $\alpha = e^{-\Delta t}$. But Δt is $1/(\text{stutter length})$ because stimulus longevity is constant no matter how temporally refined it is represented. The central result is that simulations of our model imply this theoretical relationship to be optimal for sequence learning, a result foreshadowed by the compression result in [15].

ACKNOWLEDGMENT

This work was supported by David A. Harrison III research grants to K.E.M and P.L., by NIH MH48161, MH63855 and RR15205 to WBL; and NSF NGSEIA-9974968, NPACI ASC-96-10920, and NGS ACI-0203960 to Marty Humphrey. We also thank Aaron Shon, Sean Polyn, David Sullivan, and Xiangbao Wu for their comments on early versions of the text.

REFERENCES

[1] W.B Levy, "A computational approach to hippocampal function," in *Computational modeling of learning in simple neural systems*, R. D. Hawkins and G. H. Bower, Eds. Orlando, FL: Academic Press, 1989, pp. 243 - 305.

[2] W.B Levy, X.B. Wu, and R.A. Baxter, "Unification of hippocampal function via computational/encoding considerations." in Proc. Third Workshop on Neural Networks: from Biology to High Energy Physics. *Intl. J. Neural Sys.*, vol. 6, (Supp.), (D. J. Amit, P. del

Guidice, B. Denby, E. T. Rolls and A. Treves, Eds.), Singapore:World Scientific Publishing, 1995, pp. 71-80.

[3] W.R. Holmes and W.B Levy, "Insights into associative long-term potentiation from computational models of NMDA receptor-mediated calcium influx and intracellular calcium concentration changes," *J. Neurophysiol.*, vol. 63(5), pp.1148-68, 1990.

[4] K.I. Blum and L.F. Abbott, "A model of spatial map formation in the hippocampus of the rat," *Neural Comput.*, 8, pp. 85-93. 1996.

[5] M.R. Mehta, C.A. Barnes, and B.L. McNaughton, "Experience-dependent, asymmetric expansion of hippocampal place fields," *Proc. National Acad. Sci.*, USA, 94, pp.8918-8921. 1997.

[6] M.R. Mehta, M.C. Quirk, and M.A. Wilson, "Experience-dependent, asymmetric shape of hippocampal receptive fields," *Neuron*, 25, pp.707 -715. 2000.

[7] A.P. Shon, X.B. Wu, D.W. Sullivan and W.B Levy, "Initial state randomness improves sequence learning in a model hippocampal network," *Phys. Rev. E*, vol 65, 031914. 2002.

[8] X.B. Wu, J. Tyrcha, and W.B Levy, "A neural network solution to the transverse-patterning problem depends on repetition of the input code," *Biol. Cybern.*, 79, 1998, 203-213.

[9] D. Marr, "Simple memory, a theory for archicortex," *Phil. Trans. Roy. Soc. London*, 262, 23-81, 1971

[10] E.T. Rolls, "A theory of hippocampal function in memory," *Hippocampus*, 6, pp. 601-620. 1996.

[11] W.B Levy, and O. Steward, "Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus," *Neuroscience*, 8, pp.791-797. 1983.

[12] W.B Levy, "A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks," *Hippocampus*, 6(6), pp. 579 - 591. 1996.

[13] W.B Levy, and P. Sederberg, "A neural network model of hippocampally mediated trace conditioning," in *IEEE International Conference on Neural Networks*, vol. 1, pp. 372-376. Piscataway, NJ: Institute of Electrical and Electronics Engineers. 1997.

[14] P. Rodriguez, and W.B Levy, "A model of hippocampal activity in trace conditioning: Where's the trace?" *Behav. Neurosci.* 115, 2001, 1224-1238.

[15] D.A. August, and W. B Levy, "Temporal sequence compression by an integrate-and-fire model of hippocampal area CA3." *J. Computational Neurosci.*, 6, 1999, 71-90.