

Context codes and the effect of noisy learning on a simplified hippocampal CA3 model

Xiangbao Wu, Robert A. Baxter, William B Levy

Department of Neurosurgery, Box 420, University of Virginia Health Sciences Center, Charlottesville, VA 22908, USA

Received: 17 March 1995/Accepted in revised form: 19 September 1995

Abstract. This paper investigates how noise affects a minimal computational model of the hippocampus and, in particular, region CA3. The architecture and physiology employed are consistent with the known anatomy and physiology of this region. Here, we use computer simulations to demonstrate and quantify the ability of this model to create context codes in sequential learning problems. These context codes are mediated by local context neurons which are analogous to hippocampal place-coding cells. These local context neurons endow the network with many of its problem-solving abilities. Our results show that the network encodes context on its own and then uses context to solve sequence prediction under ambiguous conditions. Noise during learning affects performance, and it also affects the development of context codes. The relationship between noise and performance in a sequence prediction is simple and corresponds to a disruption of local context neuron firing. As noise exceeds the signal, sequence completion and local context neuron firing are both lost. For the parameters investigated, extra learning trials and slower learning rates do not overcome either of the effects of noise. The results are consistent with the important role played, in this hippocampal model, by local context neurons in sequence prediction and for disambiguation across time.

1 Introduction

Early on (Amari 1972; Fukushima 1973), and more frequently in recent years, recurrently connected networks of neuron-like elements have been studied as sequence learning systems (Sompolinsky and Kanter 1986; Kleinfeld 1986; Jordan 1986; Buhmann and Schulten 1987; Coolen and Gielen 1988; Mozer 1989; Bauer and Krey 1990; Elman 1990; Reiss and Taylor 1991; Heskes and

Gielen 1992; Bartholomeus and Coolen 1992; Griniasty et al. 1993).

Concurrently with these recent studies, Levy (1985, 1989, 1994) hypothesized that the cognitive functions of the hippocampus, and its related paralimbic structures, will be unified by considering the computations performed by the hippocampus and related structures in terms of sequence learning and conditional code compression. That is, the diverse cognitive/behavioral functions that depend on the hippocampus will be inter-related by the computational architecture (sparse recurrence) and computational functions (e.g., creation of context codes from multimodal sequences) specialized to the hippocampus. Among the most important of these diverse functions are its roles supporting cortical encoding of declarative memories (Squire 1992), cognitive mapping (O'Keefe and Nadel 1978), and its hypothesized ability to mix signals and remove redundancy from cortical codes (Levy 1985, 1989; Amaral and Witter 1989; Eichenbaum and Buckingham 1991; Gluck and Myers 1993a, b).

In contrast to ideas related to memory encoding and cortical teaching, the function of the hippocampus as a cognitive mapping system has been studied extensively in the rat (e.g., O'Keefe and Nadel 1978; Eichenbaum et al. 1989; Breese et al. 1989; McNaughton and Nadel 1989; Muller and Kubie 1989). These studies propose that the hippocampus recodes and retains information about an organism's environment in such a way that this information can be used to navigate the real world in a flexible fashion (Eichenbaum et al. 1989). Our theory of hippocampally mediated cognitive mapping is simple. From codes of individual places – including distances from objects and their azimuth as used in theories by Sharp (1991), Saypoff et al. (1992), Schmajuk and Thieme (1992) and others – sensory sequences are produced via the association of individual representations over time. By learning these sequences, an organism gains the ability to navigate through its world. That is, from previously learned sequences of places, the organism can predict the paths that will carry it from one place to another. A cognitive map is learned from input sequences, and it

consists of coded representations of sequences of features (such as landmarks and other multimodal sensory cues) which prove to be useful for navigation. In rats, a cognitive map is presumably mediated by place cells (O'Keefe and Nadel 1978). Thus, by analogy, the context codes of learned sequences are hypothesized to form cognitive maps.

From one viewpoint (Levy 1989; see also Kesner and Hardy 1983; Kesner 1990, 1991 for a related viewpoint), cognitive mapping is equivalent to sequence prediction and both are related to episodic learning via learning of context. A succession of points or places an animal visits during exploration of its environment is a sequence which the animal needs to remember, generalize and recall later during physical or mental navigation of the environment. A sequence of places up to a given point in time and space also forms the context for any decision made at that point. It is with this logic, and it is by concentrating on sequence learning, that our model promises to unify hippocampal function.

When studying such a sequence prediction model, a particular pattern of neuronal activity, called local context neuron firing, is critical to good network performance. By virtue of the local context codes, our CA3-like model solves several types of sequence prediction problems when noise-free patterns are the inputs. In particular, this simple model can learn sequences in a single trial and then perform sequence completion (Minai and Levy 1993c), do jump-ahead prediction after learning a sequence (Prepscius and Levy 1994) and disambiguate two sequences that share a subsequence (Minai et al. 1994, Levy et al. 1995). This model also solves two sequence prediction problems relevant to cognitive mapping: finding shortcuts and seeking goals (Levy et al. 1995).

In the present paper, we show that the network encodes context via neural firing associated with particular subsequences and that the network can use these context codes to solve disambiguation problems. Two results growing out of this context coding are notable. First, these context codes are like place cells. Second, noisy input sequences during learning simultaneously destroy prediction and local context neuron firing.

2 The computational architecture and methods

A generally accepted, gross computational architecture of the hippocampus (e.g., Levy 1989; Eichenbaum and Buckingham 1991; Hasselmo and Schnell 1994; O'Reilly and McClelland 1994) has three parts, i.e., an input layer, a recoder inspired by the hippocampal CA3 region, and a decoder of the recoded signals inspired by the hippocampal CA1 region. This study focuses on computations performed by a CA3-like structure, and we leave for another paper detailed research on the decoder role of a CA1-like structure.

Rather than viewing CA3 as a content addressable memory (Treves and Rolls 1992), our CA3 learns and predicts sequences. Also distinctive from many models (e.g., Reiss and Taylor 1991), is that our network defines its own codes. The architecture of this CA3-like model

has four essential aspects: (1) sparse recurrent excitatory connectivity; (2) a neuronal delay of at least one time step in converting an input to an output (i-o); (3) an associative modification rule that spans at least the i-o time step; and (4) some generic feedback inhibition that narrowly, but imperfectly compared with competitive networks, bounds total activity.

The critical aspect of CA3 is its sparse random recurrent connectivity that implies an asymmetric recurrence. In most of our simulations here, the recurrent CA3 connectivity is only 5%, i.e., a neuron connects to about 25 other neurons of the 512 neurons used in the CA3 layer of our model (10% and 15% connectivity are used in our investigations elsewhere with qualitatively the same results). As is well known in the neural network field (Hopfield 1982; Cohen and Grossberg 1983), symmetry leads to stable states of network activity while an asymmetric system provides richer dynamics (Amari 1971; Minai and Levy 1993a, b, 1994; Wu and Liljenström 1994; Liljenström and Wu 1995).

In this study, the networks consist of an input layer (corresponding to layer II of the entorhinal cortex) and a sparsely connected feedback layer (CA3-like). To simplify things, single lines, x_j , from the input layer can fire a CA3 neuron. The recurrent layer consists of 512 binary (0/1) primary neurons with identical firing thresholds, θ . The neurons are interconnected via a Bernoulli process: each neuron j has a probability ($p = 0.05$) of receiving a modifiable excitatory connection from each neuron i in the recurrent layer. The presence of such a connection is indicated by the binary, $\{0, 1\}$, variable c_{ij} . Inhibition is mediated by a single interneuron that receives input from all primary neurons in the layer; this interneuron then provides an identical shunting conductance proportional to its input to all primary neurons. At time t , taking $w_{ij}(t)$ as the excitatory weight from neuron i to j , K_I as the fixed inhibitory weight from the input layer and K_R as the fixed weight for feedback inhibition, the excitation y_j of neuron j is

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

and the output of CA3 neuron j is

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

Note that an active external input always produces a firing while neurons with no active external input are not forced to the zero state and could be fired through feedback connections. A postsynaptic rule (Levy 1982), i.e.,

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

is used in our simulations, where i is input and j is output. In the past, we have used other local associative rules (Minai and Levy 1993c) and found little difference in performance.

The time span of the associative modification rule is kept to a minimum (one time step). Although such a short time is biologically unrealistic (assuming the processing

time of a neuron is 20–40 ms), we want to demonstrate the power of the network architecture for solving sequence problems that require more than one time step of memory. In fact, our own experimental results show that the time span of associativity is of the order of 50–150 ms (Levy and Steward 1983). Indeed, such a time span has been incorporated in the hippocampal model of Abbott and Blum (1995) and is central to their results. Incorporating this longer time span into our model will only increase its abilities to make predictions across longer time spans. Also like the Abbott and Blum model, our model can make step-by-step predictions in real time when the external inputs are active.

For the quantitative simulations analyzed here, the network receives sequences of 40 patterns. In contrast to the disambiguation problem, the patterns change slowly over time. They shift 1 active bit per unit of time. In terms of the noise-free prototype patterns, each external input pattern contains 8 on-bits (i.e., the firing of 8 input neurons). As a result, the Hamming distance between a pair of successive prototype patterns is 2 (for the noise-free prototype patterns). However, instead of presenting the network with deterministic sequences made up of this sequence of 40 noise-free patterns during learning, the input sequence comprises 40 randomly corrupted versions of these patterns. Corruption occurs by randomly complementing the eight input lines that are turned on externally at each time step with probability 0.1 and by randomly complementing each of the 504 external lines that are prototypically inactive at each time step with another probability (either 0.002, 0.004, 0.006, 0.008, 0.01, 0.012 or 0.02). The noise-to-signal ratio (N/S) quantifies the different noise levels. The signal has a Hamming length of 8. The Hamming distance between each corrupted input pattern and the uncorrupted input pattern at each time point is the instantaneous noise of the input. The ratio N/S is calculated as the average of the number of differing input neurons at each time step divided by the signal length.

The network is presented with the entire sequence of 40 patterns for 300 learning trials, and is allowed to modify its weights according to the postsynaptic learning rule. Before each trial, CA3 neurons are randomly excited. During recall testing, the network is randomized, then given an uncorrupted version of the first pattern of the sequence as a prompt, and then allowed to relax. To evaluate network performance after learning, we look at the end of the sequence. The average similarity (calculated as the cosine of the angle between network state vectors) of states from time step 41 to time step 60

generated in response to the single-pattern probe test is compared with the last state produced by the full sequence of 40 input patterns.

3 Results

3.1 Disambiguation of overlapping sequences

The most direct proof that the network learns context is demonstration of sequence disambiguation. Using a sequence completion task, we demonstrate the disambiguation of sequences with temporal overlaps of more than one step. The basic premise of the disambiguation shown here is that simple one-step associative learning can not solve this problem. The problem is only solvable if some neurons not directly stimulated by the input patterns are also allowed to fire. This auxiliary activity creates distinguishable, context-dependent secondary representations of the stimulus patterns and allows the correct transitions to be learned. Via these neurons, the input stimulus is mapped to a higher-dimensional space where uncertainties can be distinguished by using the extra dimensions.

Two input sequences of 20 patterns (Fig. 1), with a shared subsequence (from pattern 6 to pattern 10), are used. However, instead of presenting the network with deterministic sequences (Minai et al. 1994; Prepscius and Levy 1994), randomly corrupted versions of the two input sequences are presented during learning. Each sequence is made up of three subsequences (of lengths 5, 5 and 10), and the members of one subsequence are orthogonal to any member of another subsequence (for the noise-free prototype patterns).

To test recall, all neurons are reset to random initial values, and the system receives the first pattern associated with one of the two sequences. The network is then allowed to relax without further input for 40 time steps. Its output at the final step of the sequence can be compared with the corresponding true pattern of the fully driven sequence. From Fig. 2 it can be seen that the network predicts the end of sequence-1 rather than sequence-2 in response to pattern one of sequence-1. Not shown is the complementary behavior when pattern one of sequence-2 is given.

3.2 Context neuron firings (CNF)

Figure 3 shows what it means for the network to create its own compressed code (Levy 1989). After training, the

1A. Sequence-1 (Neurons 301-512 not shown)



1B. Sequence-2 (Neurons 301-512 not shown)



Fig. 1A, B. Two input sequences of 20 patterns, with a shared subsequence (from pattern 6 to pattern 10) are used for a disambiguation problem. Randomly corrupted versions of the two input sequences are presented during learning. Because of space limitations, here (as in other figures) only the first 300 of the 512 neurons are shown. Time is read from top to bottom, and neurons are read from left to right. A large dot indicates a neuronal firing. A Sequence-1; B sequence-2

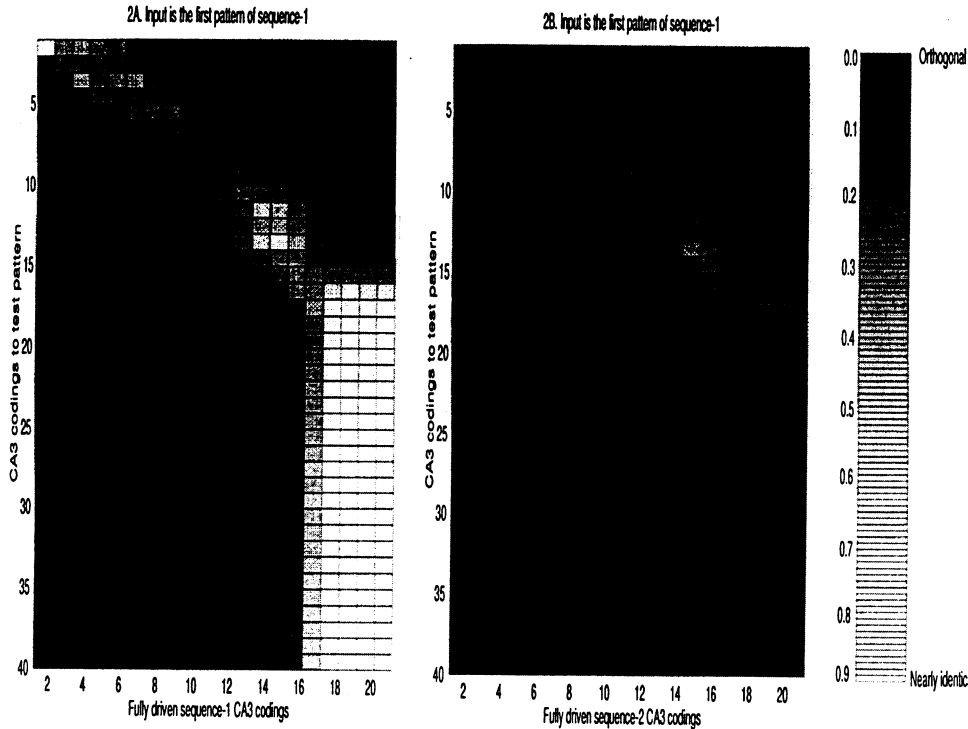
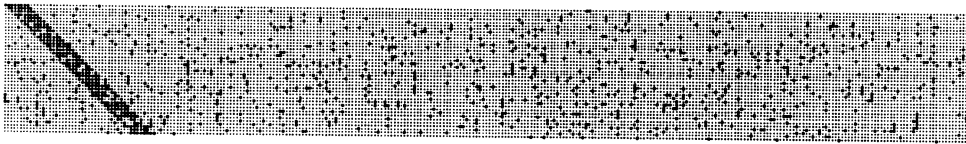


Fig. 2A, B. Test results in response to the first pattern of sequence-1. Each graph compares the similarity of successive CA3 state vectors (going from *top to bottom* of the ordinate) elicited by input pattern 1 of sequence-1 against every state vector elicited by one particular complete input sequence (going from *left to right* on the abscissa). The comparison is done by calculating the cosine of the angle between two CA3 state vectors. If two vectors are very similar, the *shade* is very light; if two vectors are nearly orthogonal, the *shade* is very dark. By comparing A with B it is clear that the network predicts the end of sequence-1 rather than sequence-2 in response to pattern 1 of sequence-1

3A. Representative CA3 activities before training (Neurons 301-512 not shown)



3B. Representative CA3 activities after training (Neurons 301-512 not shown)

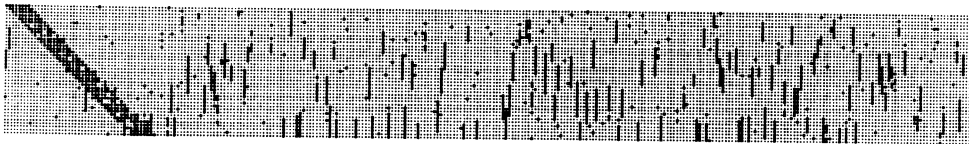


Fig. 3A, B. The development of context cell firing. **A** Neuron firing in response to an input sequence before learning. **B** Neuron firing in response to the input sequence after 300 presentations of the input. Time is read from *top to bottom*, and neurons are read from *left to right*

auxiliary neuron (Fig. 3B) activity becomes more regular in comparison with before training (Fig. 3A). Note how the CA3 encodings of successive patterns are similar to each other but still different enough to be distinguished. Note also the distinctive activity patterns of the many local context neurons. Such neurons usually switch on just once, and once on, they tend to stay on for 4–10 successive time points. Such local context neurons greatly outnumber the externally driven neurons. It is their pattern of firing that supplies the information for context-dependent prediction in a flexible fashion.

3.3 Noise effects on context codings and network performance

In our studies of disambiguation with noise-free and noisy sequences, we noted a deterioration in performance as more and more noise is added (unpublished observa-

tions). Therefore, the remainder of the results will explore the role of noise during learning. We examine both local context neuron firings (CNF) and performance in a sequence prediction problem. In particular, we study sequence prediction in the simplified setting of learning only a single sequence.

As expected, noise during learning affects the performance of the network, and it can be quite harmful to performance if N/S is too high. As noise level increases, performance decreases and eventually the noise destroys the sequence completion (Fig. 4). Note Fig. 4A is noise free, and an orderly sequence completion occurs in response to a single input pattern. In Fig. 4B, performance begins to deteriorate due to noise inputs during learning. Finally, in Fig. 4C, sequence completion fails when the noise level is too high.

These effects of noise are surprisingly resistant to slower learning. Typically, we use the same constant

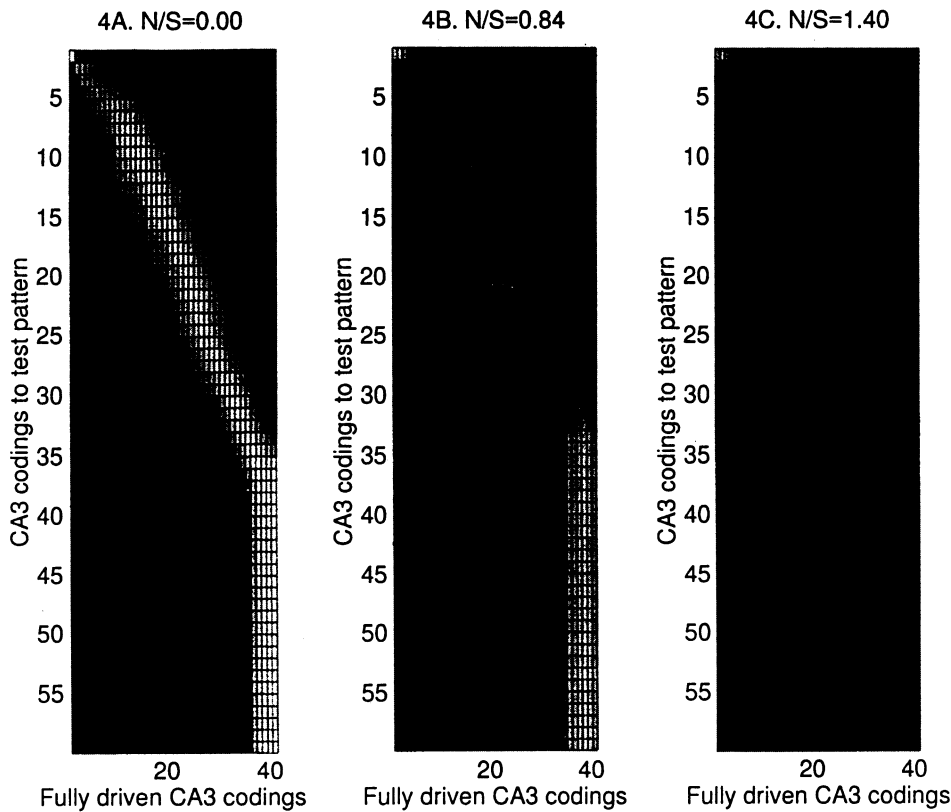


Fig. 4A–C. Noise adversely affects performance of the network. Each graph compares the similarity of successive CA3 state vectors (*ordinate*) elicited by one particular test input pattern against every state vector elicited by a complete input sequence (*abscissa*). The comparison is calculated as in Fig. 2, and the same shading scale is used. It can be seen that the network performance gets worse as the noise level increases, so that in C the network seems to fail entirely in terms of reaching the final endpoint of the sequence. A Noise-to-signal ratio (N/S) = 0.00; B N/S = 0.84; C N/S = 1.40 (see Methods for the definition of N/S and the technique that generates noisy sequences)

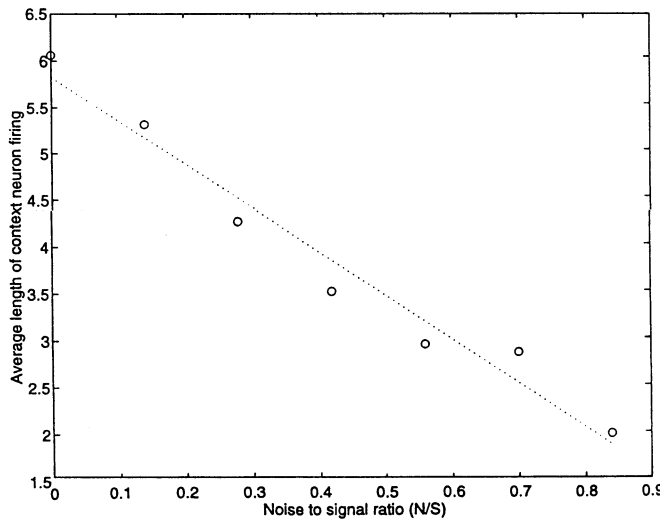


Fig. 5. Noise effects on the length of local context neuron firings (CNF). Local CNF are the recurrent activations of a neuron for sequence of time. The length of a local CNF is the number of time steps that a given local context neuron stays on. Average length of local CNF decreases almost linearly as N/S increases. The correlation coefficient of the linear fit here is about 0.98

learning rate ($\varepsilon = 0.02$) in all simulations. However, longer training time with a lower learning rate results in little difference in network performance. For example, when a learning rate $\varepsilon = 0.005$ is used during 3000 learning trials with an N/S of 0.84, the network performance is 0.64, which is very close to the value of 0.62 when using 300 learning trials with $\varepsilon = 0.02$.

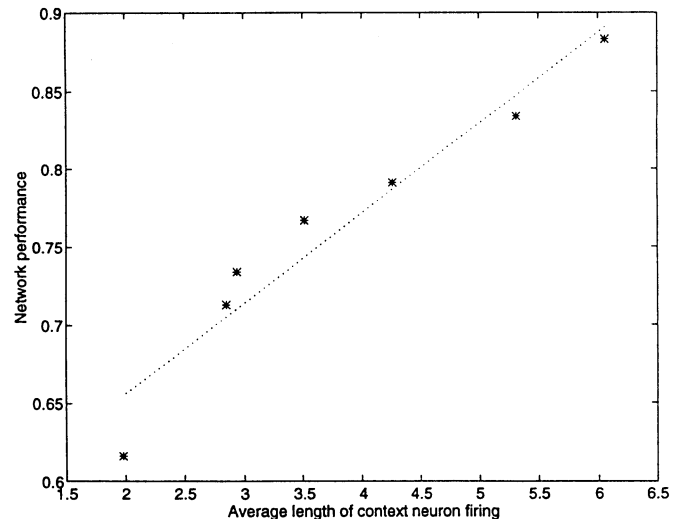


Fig. 6. Local context neuron activity predicts performance. This figure plots performance as a function of average length of local CNF. Noisy input sequences described in Fig. 3 are used during learning. Average similarity between test-elicited states, from time step 41 to time step 60, and the last state of a fully driven sequence (i.e., the last row of the last column of the similarity comparisons of Fig. 4) is plotted against average length of local CNF. The performance improves as length of local CNF increases. The correlation coefficient of the linear fit here is about 0.97, but see comments in the text. Similarity is calculated as the cosine of the angle between two state vectors

To understand why noise destroys performance, we turn to the firing patterns of context neurons. In particular, we look at the average length of local CNF, that is the number of time steps that a given local context neuron

stays on. To calculate the average length of local CNF, only neurons 48 through 512 are counted; the externally driven neurons are ignored. The average length of local CNF decreases almost linearly as N/S increases (Fig. 5).

Figure 6 shows how performance varies as a function of average length of local CNF. Note that network performance seems to improve linearly as the average length of local CNF increases. This apparent linear relationship between performance and local CNF is somewhat misleading. At the highest noise levels (e.g., Fig. 4C; N/S = 1.40), we know that a small leftward extrapolation of the plot in Fig. 6 will go to near zero; thus the relationship is not really linear. That is, as the average length of local CNF approaches 1, performance must go nearly to 0. Note that such a precipitous decrease is already present at length 2. Still, this nonlinearity does not stop us from concluding that local CNF is a crucial determinant of sequence prediction performance by the model.

4 Discussion

We have noted previously (e.g., Levy 1994) that context codes are important to the computational abilities of our hippocampal model. By studying the effect of noise, we have quantified a critical relationship between context cell firing and sequence prediction. To our way of thinking sequence learning itself, including the ability to achieve sequence completion and fast-time prediction, is part of the solution to the cognitive mapping problem. As an animal moves through its environment, a sequence is created. Such a sequence is then recoded in a compressed fashion (Levy 1989).

The shared neurons of the compressed code are the neural representation of context. We hypothesize that the hippocampus creates such compressed context codes in order to guide modification of the cerebral cortex. These context codes imply the shorter, bridging pathways between distant cortical regions that need to be associated when short-term memory is converted into a long-term store. On the other hand, in viewing sequence learning as part of cognitive mapping, we need only realize that an organism exists in time. Any series of inputs are thus elements of a sequence. Such elements encode the animal's input environment, including both the physical space in which the animal exists and any internally generated predictions about where it is or where it will be. Because the environment includes both local cues and distal cues, neighboring events in time will tend to be coded similarly by virtue of being neighbors in physical space. Patterns that tend to be sequential will be represented by an even greater coding similarity. The network, by virtue of its architecture and synaptic modification rules, will tend to compress this coding of sequences to one of greater similarity when prediction across time is valid (Levy 1989) and similar end points occur. This local context coding method has the advantage of allowing an animal to do rapid, approximate inference, i.e., the information can be read out at a range of speeds by adjusting the feedback inhibition of the network (Prepcius and Levy 1994).

By using context codes, many sequences can be successfully stored and retrieved. In the disambiguation problem shown in this paper, for instance, two sequences are stored. In a result not shown here, the network successfully learns and recalls 12 sequences of length 8. As long as the summed length of such sequences does not exceed the network's, single-sequence memory capacity, the model can learn more sequences.

Importantly, the network not only *learns* context but can also *use* context. The most direct proof that it is using context is the disambiguation problem solved in Fig. 2. Predicting the correct ending pattern implies differential representation of the shared (i.e., ambiguous) subsequence. That is, the shared (ambiguous) subsequence is differentially represented depending on the information given several patterns back in time. In other words, knowledge of the shared subsequence alone without the preceding patterns is totally – between the two choices – ambiguous for what follows. It is the earlier inputs that provide the context necessary for a correct prediction. Therefore, the network has preserved some information of pattern 1 throughout the representation of the shared subsequence. This preserved information of the recent past is, by definition, context.

The quantitative findings in this paper concern noise during learning and the disruption of prediction performance and local CNF. Noise is unavoidable in biological systems and is something nature has learned to live with. Our results show that noise adversely affects both the length of context neuron firing (Fig. 5) and performance (Fig. 6). Importantly, if the noise is too strong, it will destroy a sequence completion (Fig. 4C) because it shortens context life (Fig. 5).

In sum, we have preliminary evidence for a critical, CA3-like architecture/physiology which has great computational power relative to its simplicity. It can encode context which leads to a sequence prediction capability. Moreover, there is a strong relationship between local CNF and network performance.

Acknowledgements. This work was supported by NIH MH48161, MH00622, and EPRI RP8030-08 to W.B.L., and by the Department of Neurosurgery, Dr. John A. Jane, Chairman.

References

- Abbott LF, Blum KI (1995) Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex*, in press
- Amaral DG, Witter MP (1989) The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience* 31:571–591
- Amari S (1971) Characteristics of randomly connected threshold-element networks and network systems. *Proc IEEE* 59:35–47
- Amari S (1972) Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Trans Comput* 21:1197–1206
- Bartholomeus M, Coolen ACC (1992) Sequences of smoothly correlated patterns in neural networks with random transmission delays. *Biol Cybern* 67:285–290
- Bauer K, Krey U (1990) On learning and recognition of temporal sequences of correlated patterns. *Physica B: Condensed Matter* 79:461–475

- Breese CR, Hampson RE, Deadwyler SA (1989) Hippocampal place cells: stereotypy and plasticity. *J Neurosci* 9:1097-1111
- Buhmann J, Schulten K (1987) Noise-driven temporal association in neural networks. *Europhys Lett* 4:1205-1209
- Cohen MA, Grossberg S (1983) Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Trans Systems Man Cybern* 8:815-826
- Coolen ACC, Gielen CCAM (1988) Delays in neural networks. *Europhys Lett* 7:281-285
- Eichenbaum H, Buckingham J (1991) Studies on hippocampal processing: experiment, theory, and model. In: Gabriel M, Moore J (eds) *Neurocomputation and learning: foundations of adaptive networks*. MIT Press, Cambridge, Mass., pp 171-231
- Eichenbaum H, Wiener SI, Shapiro ML, Cohen NJ (1989) The organization of spatial coding in the hippocampus: a study of neural ensemble activity. *J Neurosci* 9:2764-2775
- Elman JL (1990) Finding structure in time. *Cognit Sci* 14:179-211
- Fukushima K (1973) A model of associative memory in the brain. *Kybernetik* 12:58-63
- Gluck MA, Myers CE (1993a) Adaptive stimulus representations: a computational theory of hippocampal-region function. In: Hanson SJ, Cowan JD, Giles CL (eds) *Advances in neural information processing systems 5*. Morgan Kaufmann, San Mateo, Calif., pp 937-944
- Gluck MA, Myers CE (1993b) Hippocampal mediation of stimulus representation: a computational theory. *Hippocampus* 3:491-516
- Griniasty M, Tsodyks MV, Amit DJ (1993) Conversion of temporal correlations between stimuli to spatial correlations between attractors. *Neural Comput* 5:1-17
- Hasselmo ME, Schnell E (1994) Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA1: computational modeling and brain slice physiology. *J Neurosci* 14:3898-3914
- Heskes TM, Gielen S (1992) Retrieval of pattern sequences at variable speeds in a neural network with delays. *Neural Networks* 5:145-152
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci USA* 79:2554-2558
- Jordan MI (1986) Attractor dynamics and parallelism in a connectionist sequential machine. In: *Proceedings of the eighth conference of the Cognitive Science Society*. Erlbaum, Hillsdale, NJ, pp 531-546
- Kesner RP (1990) Learning and memory in rats with an emphasis on the role of the hippocampal formation. In: Kesner RP, Olton DS (eds) *Neurobiology of comparative cognition*. Erlbaum, Hillsdale, NJ, pp 179-204
- Kesner RP (1991) The role of the hippocampus within an attribute model of memory. *Hippocampus* 1:279-282
- Kesner RP, Hardy JD (1983) Long-term memory for contextual attributes: dissociation of amygdala and hippocampus. *Behav Brain Res* 8:139-149
- Kleinfeld D (1986) Sequential state generation by model neural networks. *Proc Natl Acad Sci USA* 83:9469-9473
- Levy WB (1982) Associative encoding at synapses. In: *Proceedings of the fourth annual conference of the Cognitive Science Society*, pp 135-136
- Levy WB (1985) An information/computation theory of hippocampal function. *Soc Neurosci Abstr* 11:493
- Levy WB (1989) A computational approach to hippocampal function. In: Hawkins RD, Bower GH (eds) *Computational models of learning in simple neural systems*. Academic Press, New York, pp 243-305
- Levy WB (1994) Unification of hippocampal function via computational considerations. *World Congress on Neural Networks IV*:661-666
- Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience* 8:791-797
- Levy WB, Wu XB, Baxter RA (1995) Unification of hippocampal functions via computational/encoding considerations. *Int J Neural Syst, Supplement*: 71-80
- Liljenström H, Wu XB (1995) Noise-enhanced performance in a cortical associative memory model. *Int J Neural Syst* 6:19-29
- McNaughton BL, Nadel L (1989) Hebb-Marr networks and the neurobiological representation of action in space. In: Gluck MA, Rumelhart D (eds) *Neuroscience and connectionist theory*. Erlbaum, Hillsdale, NJ, pp 1-63
- Minai AA, Levy WB (1993a) The dynamics of sparse random networks. *Biol Cybern* 70:177-187
- Minai AA, Levy WB (1993b) Predicting complex behavior in sparse asymmetric networks. In: Giles CL, Hanson SJ, Cowen JD (eds) *Advances in neural information processing systems*. Morgan Kaufmann, San Mateo, Calif., pp 556-563
- Minai AA, Levy WB (1993c) Sequence learning in a single trial. *World Congress on Neural Networks II*:505-508
- Minai AA, Levy WB (1994) Setting the activity level in sparse random networks. *Neural Comput* 6:85-99
- Minai AA, Barrows G, Levy WB (1994) Disambiguation of pattern sequences with recurrent networks. *World Congress on Neural Networks IV*:176-181
- Mozier MC (1989) A focused backpropagation algorithm for temporal pattern recognition. *Complex Syst* 3:349-381
- Muller RU, Kubie JL (1989) The firing of hippocampal place cells predicts the future position of freely moving rats. *J Neurosci* 9:4101-4110
- O'Keefe J, Nadel L (1978) *The hippocampus as a cognitive map*. Oxford University Press, Oxford
- O'Reilly RC, McClelland JL (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a tradeoff. *Hippocampus* 4:661-682
- Prepscius C, Levy WB (1994) Sequence prediction and cognitive mapping by a biologically plausible neural network. *World Congress on Neural Networks IV*:164-169
- Reiss M, Taylor JG (1991) Storing temporal sequences. *Neural Networks* 4:773-787
- Saypoff R, Muller RU, Kubie JL (1992) How place cells connected by Hebbian synapses can solve spatial problems. *Soc Neurosci Abstr* 2:1211
- Schmajuk N, Thieme A (1992) Purposive behavior and cognitive mapping: a neural network model. *Biol Cybern* 67:165-174
- Sharp PE (1991) Computer simulation of hippocampal place cells. *Psychobiology* 19:103-115
- Sompolinsky H, Kanter I (1986) Temporal association in asymmetric neural networks. *Phys Rev Lett* 57:2861-2864
- Squire LR (1992) Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99:195-231
- Treves A, Rolls ET (1992) Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2:189-200
- Wu XB, Liljenström H (1994) Regulating the nonlinear dynamics of olfactory cortex. *Network: Comput Neural Syst* 5:47-60