

## Predicting the Distribution of Synaptic Strengths and Cell Firing Correlations in a Self-Organizing, Sequence Prediction Model

**Asohan Amarasingham**

*Echols Scholars Program, College of Arts and Sciences, and Department of Neurological Surgery, University of Virginia, Charlottesville, VA 22908, U.S.A.*

**William B. Levy**

*Department of Neurological Surgery and Department of Psychology, University of Virginia, Charlottesville, VA 22908, U.S.A.*

This article investigates the synaptic weight distribution of a self-supervised, sparse, and randomly connected recurrent network inspired by hippocampal region CA3. This network solves nontrivial sequence prediction problems by creating, on a neuron-by-neuron basis, special patterns of cell firing called local context units. These specialized patterns of cell firing—possibly an analog of hippocampal place cells—allow accurate prediction of the statistical distribution of synaptic weights, and this distribution is not at all gaussian. Aside from the majority of synapses that are, at least functionally, lost due to synaptic depression, the distribution is approximately uniform. Unexpectedly, this result is relatively independent of the input environment, and the uniform distribution of synaptic weights can be approximately parameterized based solely on the average activity level. Next, the results are generalized to other cell firing types (frequency codes and stochastic firing) and place cell-like firing distributions. Finally, we note that our predictions concerning the synaptic strength distribution can be extended to the distribution of correlated cell firings. Recent published neurophysiological results are consistent with this extension.

### 1 Introduction ---

The hippocampus is arguably involved with sequence prediction and the learning of context (Blum & Abbott, 1996; Levy, 1989; Hirsh, 1974; Nadel & Willner, 1980; Kesner & Hardy, 1983; Gray, 1982). Thus we have been studying self-organizing models of context-dependent sequence prediction using a simplified anatomy and physiology inspired by the biology of the hippocampus, particularly region CA3.

Such networks can learn to perform psychologically inspired paradigms believed to belong in the province of hippocampal functions (for a summary,

see Levy, 1996). These paradigms include context-dependent problems such as disambiguation, in which two different input sequences share a common subsequence (Minai, Barrows, & Levy, 1994); goal finding without search (Levy, Wu, & Baxter, 1995; Levy, 1996); certain nonlinear problems such as transverse patterning (Levy, Wu, & Tyrcha, 1996; Wu, Tyrcha, & Levy, 1996b); as well as simpler problems such as transitive inference (Levy & Wu, 1997).

The coding, that is, cell firing patterns, that our network develops to represent sequential information is interesting. Special firing patterns identify specific, temporally contiguous subsequences within the overall sequence being learned. That is, the neurons in this model, particularly the (feedback) neurons that receive no input from the external layer sequence, learn to fire exclusively within specific temporal bounds of the sequence; we refer to these neurons as *local context* units. These local context units are hypothesized to be an analog of hippocampal place cells (Wu, Baxter, & Levy, 1996a) because place cells fire briefly as a rat walks through a place field (O'Keefe & Nadel, 1978). Moreover, these firing patterns fit the generic description of synfire chains suggested by Abeles (1991) and others, with temporally adjacent local context units serving as the synfire links in Abeles's definition.

Compared to topographically organized models of sensory cortices, our hippocampal model has random recurrent connections so that it is largely structureless. However, with learning (i.e., local, online, associative synaptic modification), some structure may appear in the distribution of synaptic weights, and this structure might well reflect the structure of the input environment. Because the modification of synaptic strengths is arguably the fundamental basis of learning, understanding how the statistical structure of the input environment is impressed onto the synaptic weight matrix is a critical preparatory step for understanding how a network's dynamics are altered as a result of interaction with the environment. Here we pursue these ideas and look for such structure by developing the interrelationship between local context unit lifetime and the steady-state statistics of the synaptic weights. The analysis leads to three unanticipated results:

1. The distribution of nonzero synaptic strengths does not converge to a Gaussian distribution but rather is essentially uniform in distribution.
2. A large majority of synapses is driven to zero strength.
3. The fraction of synapses driven to zero (and therefore the uniform distribution also) can be approximately predicted by activity levels alone. This means that the probability distribution of synaptic weights can be well approximated without specific knowledge of the input environment so long as local context codes form, or, equivalently, a

noise-free input environment does not leave its impression on the weight distribution beyond its effect on activity levels alone.

## 2 Model and Methods

---

**2.1 Computational Architecture.** Our CA3 model has been described several times (Levy & Wu, 1996; Wu et al., 1996a).

The networks consist of an input layer (corresponding to layer II of the entorhinal cortex combined with the dentate gyrus) and a sparsely connected feedback layer (CA3-like). To simplify things, single axons,  $x_j$ , from the input layer always fire a CA3 neuron. The recurrent layer consists of 1024 binary {0,1} primary neurons with identical firing thresholds,  $\theta$ . The neurons are interconnected via a Bernoulli process: each neuron  $j$  has a probability ( $c = 0.10$ ) 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 the primary neurons in the recurrent 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 neuron  $j$ ,  $K_I$  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:

$$y_j = \frac{\sum_{i=1}^N c_{ij} w_{ij} z_i(t-1)}{\sum_{i=1}^N c_{ij} w_{ij} z_i(t-1) + K_R \sum_{i=1}^N z_i(t-1) + K_I \sum_{i=1}^N x_i(t)}, \quad (2.1)$$

where  $N$  is the number of neurons (1024) in the model. 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,  $x_j$ , 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.

**2.2 The Learning Rule.** A Hebbian-type postsynaptic associative modification rule is used in our simulations and analysis (Levy & Steward, 1979, 1983; Levy, 1982). For input  $i$  and output  $j$ , this postsynaptic rule is given by

$$w_{ij}(t) = w_{ij}(t-1) + \epsilon z_j(t) [z_i(t-1) - w_{ij}(t-1)], \quad (2.2)$$

where  $w_{ij}$  is the weight of the synapse connecting neuron  $i$  to neuron  $j$ . Here  $\epsilon = 0.01$ . A similar rule is used in other models (McClelland & Goddard, 1996; Treves & Rolls, 1994; Grossberg, 1982).

**2.3 Input Sequences and Learning.** Our approach has been to combine our minimal model of CA3 with a working assumption that the hippocam-

pus receives sequential sensory information—in analogy to the sequential sensory and motor input received by a rat exploring an environment. A critical assumption here has been that this external input is sparse (Levy, 1989; Treves & Rolls, 1992, 1994); many more CA3 neurons receive input from CA3 feedback than from external sources. That is, a large number of the neurons in our model do not receive any external input and fire solely as a result of feedback input.

The nature of the temporal dependencies in the external input remains an unresolved issue. For example, it seems likely that sensory stimuli do not disappear immediately but are maintained over several time steps. Nevertheless, in order to ensure accuracy on this issue, we study a wide variety of input sequences. Interestingly, varying this aspect of the input environment turns out to have minimal impact on our results.

Nine different types of input sequences were used in the simulations. Each external input pattern contained eight on-bits out of the 1024 neurons. For eight of the sequences, there was a constant shift of  $k$  bit(s) per unit of time ( $k = 1, 2, 3, 4, 5, 6, 7,$  or  $8$ ) from one pattern to the next, producing a successive overlap length of  $8 - k$  for the input patterns. In the ninth input sequence, the overlap length of external patterns varied from one time step to the next over the range 0 to 7 so that the expected overlap is 3.5; this input sequence was included to demonstrate that these results do not depend on the shift rate being constant. Figure 1A illustrates such an input sequence.

The results of the simulations described here are for networks running near their sequence length memory capacity. Sequence length capacity ( $C$ ) was reached by running simulations of different sequence lengths via a titration-like procedure. As noted previously (Levy & Wu, 1996), the capacity is simply related to average activity,  $a$ , and average local context lifetime,  $E[\ell]$  as

$$C = \frac{E[\ell]}{a}, \quad (2.3)$$

where  $E[\ell]$  is the duration of firing (in time steps) for a local context unit, averaged over all local context units in the network.

Training consisted of presenting the input sequence to the network 350 times for the random overlap sequence and 300 times for all the other sequences. Testing consisted of randomizing network activity, giving the first pattern of the learned sequence and then allowing the network to run on its own. The network's states are decoded by comparing its states at the end of training to the states during testing (Levy et al., 1995; Wu et al., 1996a; Levy & Wu, 1996). A sequence is learned if the network produces a minimum of 75% ordered recall of the recoded sequence during testing. We enforced robustness by also requiring that this definition be satisfied for at least four out of five randomly constructed networks with identical parameters. The reported data are averages of five successful simulations.

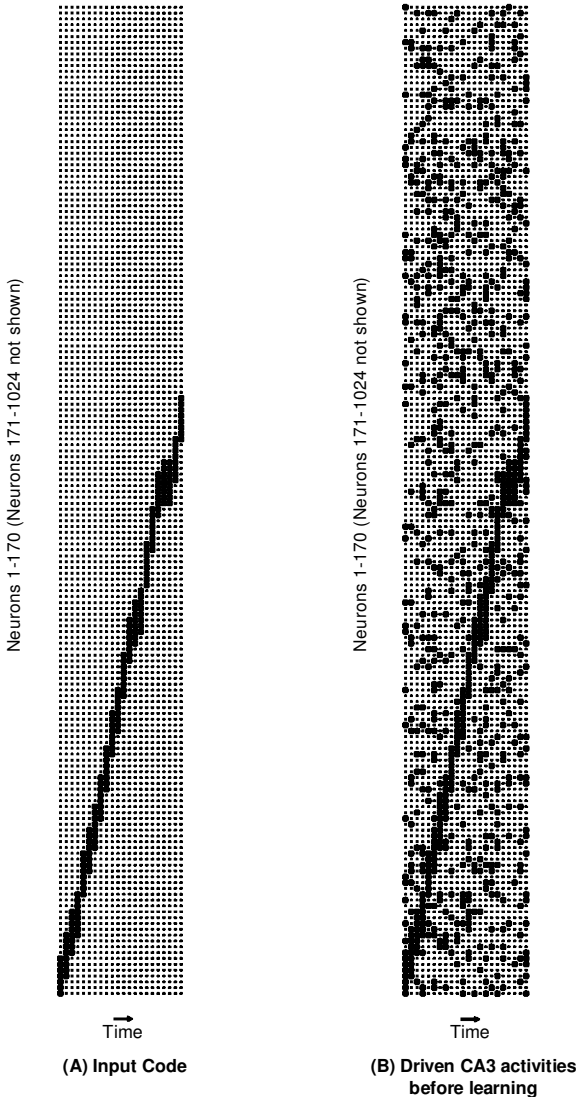


Figure 1: Typical input (A) and firing patterns (B) from a simulation of 1024 neurons before learning (only neurons 1–170 are shown). The input type is a random overlap input sequence. Large dots indicate firing (i.e.,  $z_i(t) = 1$ ); small dots indicate no firing ( $z_i(t) = 0$ ). (A) The external input code for a random overlap input sequence described in section 2. In this example, the step-to-step input overlap uniformly varies between 0 and 7. (B) Network activities when the external input sequence is presented, before learning. Notice the contrast between the firing patterns before learning, here, and the firing patterns after learning, in Figure 2A.

For the calculations used later, a small correction is made for the unused neurons that do not fire in the network's final coding. The presence of unused neurons in a steady-state coding modifies the synaptic weight prediction. We discarded these unused neurons (and therefore their synapses). Discarding unused neurons seems biologically valid (for references, see Voydovic, 1996).

### 3 Results

---

The results show that the network's steady-state synaptic weight distribution essentially depends on the distribution of local context lifetimes and the length of the input sequence. We quantify this dependence by making certain basic assumptions about the statistical properties of the model's context neurons and then confirm the suitability of these assumptions with simulations.

**3.1 Properties of the Coding Scheme—Local Context Units.** Figure 2A illustrates the recoding of the external input sequence (see Figure 1A), that the network created. Critical to the network's problem-solving abilities are the neurons that recognize a subsequence of the full sequence. We call these neurons *local context units*, and we hypothesize that they are analogous to hippocampal place cells (Levy et al., 1995). Specifically, neuron  $i$  is a local context unit if and only if there exist points in time,  $\alpha_i$  (context unit start point) and  $\beta_i$  (context unit end point), such that  $z_i(t) = 1$  if  $\alpha_i \leq t \leq \beta_i$ , and  $z_i(t) = 0$  otherwise. Consequently, we characterize the temporal length, or lifetime  $\ell_i$ , of a local context unit  $i$  as  $\beta_i - \alpha_i + 1$ , the number of time steps at which neuron  $i$  fires.

We know that local context units are crucial because we have found architectural and environmental conditions that preclude context unit formation (including high noise levels, orthogonal input sequences, and modifications in the parameters determining activity). These conditions destroy problem-solving ability just as they destroy the local context units (Wu et al., 1996a) in the context-dependent problems mentioned in Section 1. Because the network does not perform well under such conditions and also because context units are analogous to hippocampal place cells, which we know to exist, we have chosen to focus on situations in which context units do form. All of the theoretical assumptions are summarized in Table 1.

For the purposes of the theory developed here, we will assume that local context units are uniformly distributed over the sequence length; that is, the distributions of the  $\alpha_i$ 's and  $\beta_i$ 's are not overly concentrated in one region of the sequence as opposed to another. Finally, we assume that the network's coding converges (stochastically) if the activity converges (stochastically). That is, eventually the neural firing patterns reach a steady state.

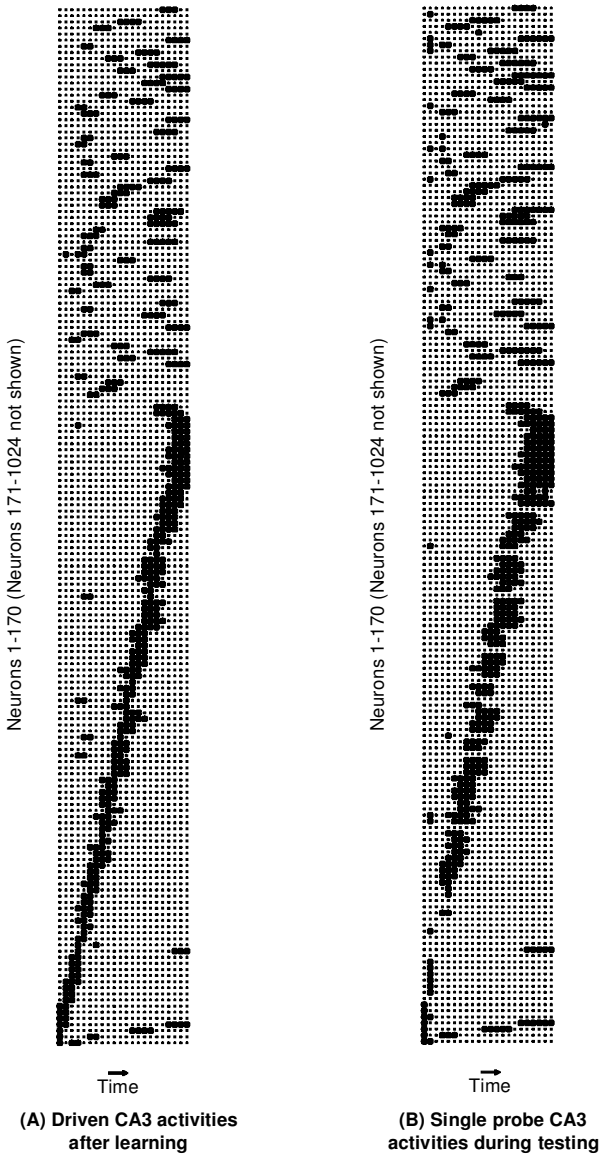


Figure 2: Typical firing patterns from a simulation of 1024 neurons after learning (only neurons 1–170 are shown). This is a continuation of the same simulation as in Figure 1. (A) Network activities when the external input sequence is presented, after learning. Notice how different the firing patterns are from Figure 1B, before learning. (B) Network activities during sequence completion testing; only the first pattern of the external input sequence is given as a probe. Notice the similarity of the firing patterns to those in Figure 2A.

Table 1: Assumptions and Approximations Used to Develop the Theoretical Arguments.

---

Assumptions

- All neurons fire as local context units. (Section 3.3.1)
- Convergence of the synaptic modification rule. (Section 3.3.2)

Approximations

- Local context units are independent of connectivity. (Section 3.3.3)
  - Local context unit starting points are uniformly distributed over the sequence length. (Section 3.3.4)
  - End effects are negligible. (Section 3.3.5)
- 

### 3.2 Relating Local Context Units and Synaptic Weights.

*3.2.1 Weight modification and neuronal firing.* As a consequence of the Hebbian-type associative synaptic modification rule (equation 2.2), synaptic weights reflect the temporal correlation between neuronal firing patterns. The more synchronous the firing activity between two neurons sharing a synapse, the greater the strength of that synapse. (This synchrony is actually displaced one time step due to the one-time-step delay in the learning rule. However, for simplicity, we will use the term synchrony forthwith, while actually meaning synchrony displaced one time step.)

Because we are assuming that neural firing patterns reach a steady state, we can consider  $z_i(t-1)$  and  $z_j(t)$  as stationary, though not necessarily independent, random variables influencing a given synaptic weight  $w_{ij}$ , via the associative modification rule (see equation 2.2). Since  $z(t)$  is  $\{0,1\}$  valued, modification occurs only when the postsynaptic neuron is active, and thus the synaptic weight becomes a running average of the conditional probability distribution  $P(z_i(t-1) | z_j(t) = 1)$ . Summing the weight modification rule recursively for this conditional case and averaging the sum as  $t \rightarrow \infty$  gives us the convergence value of the synaptic weight:

$$w_{ij} = \frac{P(z_j(t) = 1, z_i(t-1) = 1)}{P(z_j(t) = 1)}. \quad (3.1)$$

Therefore, assuming a steady state is reached, the steady-state synaptic strength is equivalent to the ratio of the number of time steps when the pre- ( $i$ ) and postsynaptic ( $j$ ) neurons are synchronously active to the number of time steps when the postsynaptic neuron is active. Cast in terms of probability, this gives:

$$P(w_{ij} = m) = P\left(\frac{P(z_j(t) = 1, z_i(t-1) = 1)}{P(z_j(t) = 1)} = m\right). \quad (3.2)$$

The essence of our idea is to calculate the two probabilities in equation 3.5,  $P(z_j(t) = 1, z_i(t-1) = 1)$  and  $P(z_j(t) = 1)$ , in terms of the temporal specifications  $(\alpha, \beta)$  of the context units  $i$  and  $j$  and the basic assumptions outlined in the previous section.

Generally the intuition is this: consider a presynaptic context unit with fixed start point and end point,  $(\alpha_i, \beta_i)$ . For each possible position  $(\alpha_j, \beta_j)$  of the postsynaptic context unit, the probabilities  $P(z_i(t-1), z_j(t))$  and  $P(z_j(t))$  are determined, and so, therefore, is their ratio (see equation 3.2). It follows that  $P(w_{ij} = m) = \sum_{\alpha_i, \beta_i, \alpha_j, \beta_j} P(w_{ij} = m \mid \alpha_i, \beta_i, \alpha_j, \beta_j) P(\alpha_i, \beta_i, \alpha_j, \beta_j)$ . The computational difficulties such a calculation involves are eased by making several simplifying assumptions. In the end, we are able to produce a closed-form solution to the conditional probability  $P(w_{ij} = m \mid \ell_i, \ell_j)$ , and then produce a numerical procedure for generating  $P(w_{ij})$ . These ideas are made more explicit in this and the following section.

The temporal correlation between pre- and postsynaptic firing patterns amounts to the temporal overlap between pre- and postsynaptic context units. This follows from the definition of a local context unit:  $z_i(t) = 1$  if and only if  $\alpha_i \leq t \leq \beta_i$ . Thus the event  $\{z_j(t) = 1\} \cap \{z_i(t-1) = 1\}$  is equivalent to the event  $\{\alpha_i \leq t-1 \leq \beta_i\} \cap \{\alpha_j \leq t \leq \beta_j\}$ . Simplifying this to the equivalent  $\{\max\{\alpha_i + 1, \alpha_j\} \leq t \leq \min\{\beta_i + 1, \beta_j\}\}$ , we define  $\text{overlap}_{ij}$ :

$$\text{overlap}_{ij} = \max \begin{cases} \min\{\beta_i + 2, \beta_j + 1\} - \max\{\alpha_i + 1, \alpha_j\} \\ 0 \end{cases} \quad (3.3)$$

so that  $\text{overlap}_{ij}$  is 0 if this difference is less than zero. Thus  $\text{overlap}_{ij}$  is the number of time steps,  $t$ , that satisfies the event  $\{z_j(t) = 1\} \cap \{z_i(t-1) = 1\}$ . Setting  $S$  as the sequence length (measured in time steps), it follows that

$$P(z_j(t) = 1, z_i(t-1) = 1) = \text{overlap}_{ij}/S. \quad (3.4)$$

The number of time steps satisfying  $\{z_j(t) = 1\}$  follows from the definition of a context unit  $j$ . Recalling the definition  $\ell_j = \beta_j - \alpha_j + 1$ ,

$$P(z_j(t) = 1) = \ell_j/S. \quad (3.5)$$

Finally, combining equations 3.4 and 3.5 via 3.2, we have

$$P(w_{ij} = m) = P\left(P\left(\frac{\text{overlap}_{ij}}{\ell_j}\right) = m\right). \quad (3.6)$$

*3.2.2 Predicting the synaptic weight distribution from a context lifetime distribution.* The next step is to formulate the conditional probabilities

$P(\text{overlap}_{ij} \mid \ell_i, \ell_j)$ . This is done by rewriting the definition of  $\text{overlap}_{ij}$  (see equation 3.3) in terms of  $\ell$ ,

$$\text{overlap}_{ij} = \max\{\min\{\alpha_i + \ell_i + 1, \alpha_j + \ell_j\} - \max\{\alpha_i + 1, \alpha_j\}, 0\}, \quad (3.7)$$

and partitioning the context lifetime probability space into three cases:  $\ell_i = \ell_j$ ,  $\ell_i > \ell_j$ , and  $\ell_i < \ell_j$ . Within these partitions, the conditional probability  $P(w_{ij} \mid \ell_i, \ell_j)$  can be simplified by making two assumptions:  $\alpha$  is uniformly distributed, and  $\alpha_i$  and  $\alpha_j$  are independent for all connected pairs  $i, j$ . In addition, to avoid considering start-up and end effects on the weight distribution, we apply the approximation to the effect that the context units are not close enough to the beginning or end of the sequence to affect the overlap probabilities (Approximation 3; see Table 1).

The calculation is made explicit for Case I ( $\ell_i = \ell_j$ ), and is summarized for  $\ell_i \neq \ell_j$  in the Appendix. The results are:

$$P\left(w_{ij} = \frac{x}{\ell_j} \mid \ell_i, \ell_j\right) = \begin{cases} \frac{s-2\ell_j-\ell_i+2}{s-\ell_j+1} & , \quad x = 0 \\ \frac{2}{s-\ell_j+1} & , \quad 0 < x < \min\{\ell_i, \ell_j\} \\ \frac{|\ell_i-\ell_j|+1}{s-\ell_j+1} & , \quad x = \min\{\ell_i, \ell_j\} \end{cases} \quad (3.8)$$

The partitioned conditional probabilities described here completely define the conditional weight distribution,  $P(w_{ij} \mid \ell_i, \ell_j)$ . Thus combining this conditional distribution with the empirical context lifetime distribution,  $P(\ell)$ , derived from network simulations, yields the distribution  $P(w_{ij} = x)$ . Using

$$P(w_{ij} = x) = \sum_y \sum_z P(w = x \mid \ell_i = y, \ell_j = z)P(\ell_i = y, \ell_j = z), \quad (3.9)$$

we then apply an additional assumption that presynaptic context length and postsynaptic context length are independent:

$$P(w_{ij} = x) = \sum_y \sum_z P(w = x \mid \ell_i = y, \ell_j = z)P(\ell = y)P(\ell = z). \quad (3.10)$$

In this way we have the complete distribution of weights based on knowledge of the distribution of local context unit lifetimes,  $P(\ell)$ .

**3.2.3 Simplifications and comparison of theory to simulations.** If predictions like those generated here have any hope of being tested in a biological experiment, we must relax the requirement for knowledge of the distribution

$P(\ell)$ . A simpler theoretical perspective arises by assuming that all local context units are of the same length,  $E[\ell]$ . Thus it follows from equations 3.8 and 3.10:

$$P(w_{ij} = 0) = \frac{S - 3E[\ell] + 2}{S - E[\ell] + 1} \quad (3.11a)$$

$$P(w_{ij} = x/E[\ell]) = \frac{2}{S - E[\ell] + 1}, \quad \text{for } 0 < x < E[\ell] \quad (3.11b)$$

$$P(w_{ij} = 1) = \frac{1}{S - E[\ell] + 1} \quad (3.11c)$$

The qualitative conclusions we might draw from equation 3.11a–c are the same as those for equation 3.8. The zero-weight proportion (see equation 3.11a) would dominate in the biologically relevant case of low activity because  $P(w_{ij} = 0) \uparrow 1$  as  $E[\ell]/S \rightarrow 0$ , and  $E[\ell]/S$  is a measure of neuronal activity. For  $w_{ij} > 0$ , the distribution will be uniform with the exception of a disproportion at  $w_{ij} = 1$ . Therefore, to further simplify the weight distribution, we calculate the  $P(w_{ij} = 0)$  from equation 3.11a and then simply project a uniform distribution for the remaining cases,  $w_{ij} > 0$ .

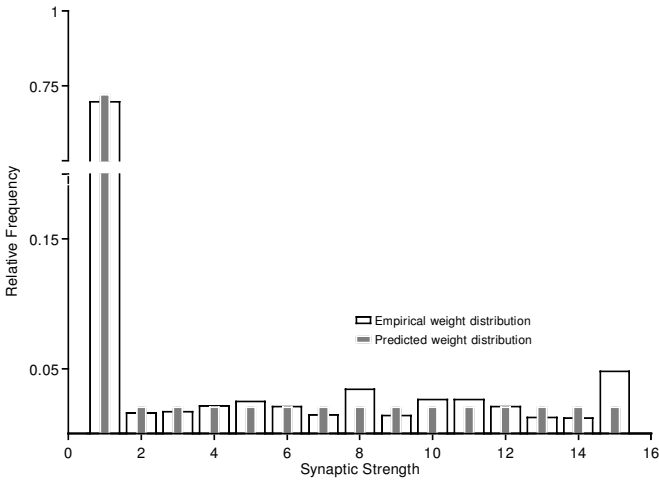
Figure 3 compares these predicted distributions to the average distribution obtained from empirical results for two input sequences. The prediction of  $P(w_{ij} = 0)$  is remarkably good in these histograms and across all nine input types. The uniform distribution of the simplified theory also does very well except at  $P(w_{ij} = 1)$ . Here the prediction is not accurate, with the largest bin being underestimated by about 50%. The error here, in a situation in which local context units are perfectly correlated, stems from the assumption that they are uncorrelated. Fortunately this assumption seems to cause very little inaccuracy elsewhere.

The impression of accuracy conveyed by these histogram comparisons is found for all nine input sequences (see Table 2, which quantifies the error in these comparisons by measuring histogram bin differences between theory and simulation).

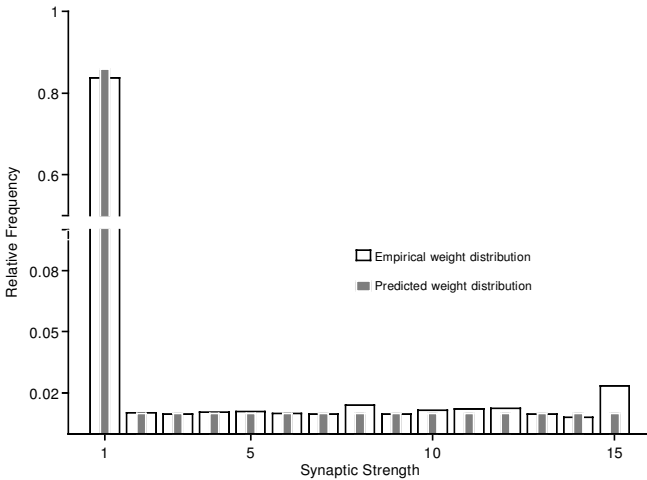
**3.2.4 Zero-Weight Approximations from Average Activity.** For a network running at sequential memory capacity, equations 3.11a and 2.3 imply a relationship that is approximately

$$P(w_{ij} = 0) \sim 1 - 2a. \quad (3.12)$$

The crudeness of this approximation depends on activity and capacity, with accuracy increasing both as the network approaches capacity (as  $S \uparrow C$ ) and as activity decreases. In the case of the overlap four-input sequence, the  $1 - 2a$



(a)



(b)

Figure 3: Synaptic weight histograms comparing simulations and theoretical predictions. The predicted weight distributions were derived by using equation 3.12 to predict  $P(w_{ij} = 0)$  and projecting a uniform distribution across the rest of the weight space. The weight space was divided into 15 bins:  $[0,1/15)$ ,  $[1/15,2/15)$ ,  $[2/15,3/15)$ ,  $\dots$ ,  $[13/15,14/15)$ ,  $[14/15,1]$ . In both A and B, the abscissa is broken and the scale changes. (A) Empirical and theoretical weight distributions after learning of the random overlap input sequence described in Figure 1A. Distributions are averages over five runs. (B) Empirical and theoretical weight distributions after learning of an overlap four external input sequence. Distributions are averages over five runs.

Table 2: Prediction Errors, Organized in Terms of the Overlap of the External Input Sequence.

Overlap	Predicted from $P(\ell)$ , Equation (3.10)		Predicted from $E[\ell]$ , Equation (3.11a)	
	Error	Var(error)	Error	Var(error)
0	1.32e-2	6.22e-4	1.36e-2	5.32e-4
1	1.40e-2	6.49e-4	1.34e-2	6.50e-4
2	1.30e-2	4.89e-4	1.16e-2	5.67e-4
3	1.02e-2	1.84e-4	7.3e-3	1.96e-4
4	7.3e-3	6.09e-5	3.7e-3	4.42e-5
5	5.5e-3	5.49e-5	3.7e-3	4.38e-5
6	5.1e-3	6.22e-5	4.4e-3	4.91e-5
7	5.8e-3	8.04e-5	4.4e-3	6.67e-5
Random	1.34e-2	8.30e-5	8.5e-3	7.16e-5

Note: The weight space,  $[0,1]$ , was divided into 15 bins, as in Figure 3. In each simulation, the error associated with bin  $k$  was measured as the absolute value of the difference between the predicted (fraction)  $P(w_{ij} \in \text{bin}_k)$  and the observed count  $(1/N^2c) \sum I(w_{ij} \in \text{bin}_k)$ , where  $N^2c$  is the number of synapses in the network. The reported error is the average value of this error measure over all bins in five simulations of each input sequence. Var(error) is the variance of the error measure over the same set.

approximation was within 1% of the prediction generated by equation 3.11a and within 3% of the empirical value for all five simulations.

For sequences of length less than capacity ( $S < C$ ), this approximation can be corrected for the total fraction of used neurons,  $u$ , via the correction  $S = uE[\ell]/a$ . This gives

$$P(w_{ij} = 0) \sim 1 - \frac{2a}{u} - \frac{2a^2}{u^2 - au},$$

which can be approximated as

$$P(w_{ij} = 0) \sim 1 - 2\frac{a}{u}, \quad (3.13)$$

for small values of the ratio  $a/u$ .

**3.3 Validity of the Theoretical Assumptions.** The central result of this article concerns the distribution of synaptic weights and is best demonstrated by the empirical results. The theory was developed in order to examine the validity of this result within the broader context of sequence-predicting neural models. A list of the assumptions used for its development is provided in Table 1.

*3.3.1 Assumption 1: Local context unit formation.* Perhaps the most critical assumption is that coding neurons behave as local context units; they all fire

Table 3: Simulation Statistics.

Overlap	Sequence Length	Activity <sup>a</sup>	$E[\ell]$	Unused Neurons <sup>b</sup>	Multiple Firing <sup>c</sup>
0	20	0.051	1.01	144	155
1	24	0.048	1.21	153	105
2	33	0.049	1.59	112	121
3	47	0.064	3.07	64	45
4	57	0.061	3.60	59	15
5	90	0.055	5.11	45	15
6	110	0.050	5.85	63	6
7	160	0.054	9.49	113	15
Random	22	0.138	2.71	77	15

Note: The data points represent averages over five simulations per input sequence.

<sup>a</sup>Defined as  $\frac{1}{NS} \sum_{i,t} z_i(t)$ , when the external input sequence is presented, after learning.  $E[\ell]$  is defined per the text.

<sup>b</sup>A count of the neurons that do not fire in the final coding.

<sup>c</sup>A count of the neurons that do not satisfy the local context unit definition (i.e., fire at different, noncontiguous time points in the sequence) in the final coding.

exclusively within specific boundaries within the sequence. The empirical results presented here certainly justify such an assumption (see Table 3, particularly the statistics concerning multiple firing neurons), but we believe that the formation of local context units is important in a broader sense, particularly with regard to hippocampal function.

Experimental support for local context units comes from recordings of place cells in the hippocampus of rats that have learned a spatial environment (O’Keefe & Nadel, 1978). Place cells fire when a rat traverses specific spatial locations within the learned environment; that is, these cells fire more or less exclusively within specific spatial boundaries. Moreover, because navigation over space is necessarily continuous with respect to time, it follows that the sensory information provided during navigation is itself sequential, and thus these neurons are firing exclusively within specific boundaries in a sequence. This firing pattern is precisely the definition of a local context unit and suggests an important relationship between the sequence-learning model presented here and experimental neuroscience.

A second justification for the relevance of local context units comes from their functional importance: local context units form associative bridges over time. If the computational challenge in sequence learning is to develop appropriate synaptic interactions between neurons representing arbitrary sequences of external stimuli, then a significant biological concern is that such connections may not exist in the first place. This issue is avoided in sequence-learning models, such as those of Griniasty, Tsodyks, and Amit

(1993) and Amit, Brunel, and Tsodyks (1994), in which each neuron is assumed to be connected to every other neuron. In the biologically relevant context of sparse connectivity, however, connections between neurons representing successive external patterns may not exist. By allowing recurrent neurons (which are not originally associated with any external stimulus) to play a role in encoding the sequence, however, this CA3 model demonstrates how a more reasonable connectivity can form the required temporal-to-spatial association, that is, by forming local context units.

The formation of local context units is also necessary for the model to solve several interesting (i.e., context-dependent) sequence-prediction problems above and beyond simple sequential recall (as discussed in section 1; for a review, see Levy, 1996).

*3.3.2 Assumption 2: Convergence of the learning rule.* The other important assumption is that the associative weight converges. This implies that synaptic weights take on a scaled correlation of presynaptic and postsynaptic firing,  $E[z_j(t)z_i(t-1)]/E[z_j(t)]$ . In the case of the binary neurons used here, this correlation reduces to the conditional probability that the presynaptic neuron fires given that the postsynaptic neuron fires. In a recurrent network, we can provide no guarantees; however, the success of the theory in predicting the empirical weight distributions is perhaps fundamentally a vindication that synaptic weights can take on these conditional probabilities. Whether convergence is achieved is an important issue, because several models (Griniasty et al., 1993; Amit et al., 1994; Bienenstock, 1995) have implicitly assumed this sort of convergence by directly loading correlations into the synaptic weights. Thus the success of this theory demonstrates a biologically plausible learning rule that learns these correlations.

Finally, because place cells are stable (Thompson & Best, 1989), in analogy to convergence, we have reason to believe that the assumption of convergence is a reasonable biological assumption.

*3.3.3 Approximation 1: Local context units are independent of connectivity.* This assumption is quite clearly an approximation, since, in an absolute sense, it is not true. In reality, the fact that two neurons are connected *does* influence the relative positions (starting points) of their local context units; it increases the likelihood that they are temporal neighbors. Because local context units are determined by two parameters, start point and lifetime, we express this approximation mathematically by the statements  $P(\alpha_i, \alpha_j | c_{ij} = 1) = P(\alpha_i, \alpha_j)$  and  $P(\ell_i, \ell_j | c_{ij} = 1) = P(\ell_i, \ell_j)$ . (When we generalize, in section 3.4, to firing patterns that are parameterized by center  $\mu$  and variance  $\sigma^2$ , rather than by starting point and lifetime, this assumption becomes  $P(\mu_i, \mu_j | c_{ij} = 1) = P(\mu_i, \mu_j)$  and  $P(\sigma_i^2, \sigma_j^2 | c_{ij} = 1) = P(\sigma_i^2, \sigma_j^2)$ .)

*3.3.4 Approximation 2: Local context unit starting points are uniformly distributed over the full sequence.* This is a fairly simple approximation which we expect to be essentially true for the constant-shift input sequences. However, the theoretical results hold quite well even when local context units are not uniformly distributed, as indicated by the results in the case of the random overlap input sequence. Thus, this does not appear to be a critical assumption, and, in fact, it would not be difficult to extend the theory to handle any alternative distribution, so long as the distribution is specified.

*3.3.5 Approximation 3: End effects are negligible.* This approximation arose out of the need to calculate the  $\alpha_i, \alpha_j$  overlap distribution given only that  $c_{ij} = 1$  (the two neurons are connected) and using Approximations 1 and 2 only. This calculation takes a simpler form when the negligible effects imposed by the skewing of the distribution at the boundaries of the sequence are ignored, so we assumed that the presynaptic local context unit was not at a boundary of the sequence—specifically, that  $\ell_i < \alpha_i < S - \ell_i$  when we calculated the overlap distribution.

The situation is analogous for the continuous-valued firing patterns, discussed in section 3.4. Here we place the center of a prototypical neuron's firing function at the center of the sequence and then calculate the probability distribution of weights for the synapses that input to this prototypical neuron. We then assume that all neurons have a similar distribution; that is, we assume that end effects are negligible.

**3.4 Alternative Cell Firing Models.** It is natural to wonder whether the results described so far are particular to the assumptions made about cell firing—specifically, the assumptions of discrete time and deterministic, binary output. In this section, we briefly illustrate the extension of our analysis to other cell firing distributions and demonstrate that this variation does not affect the essential result that the synaptic weight distribution is not gaussian or even centrally peaked.

*3.4.1 Continuous firing.* The natural continuous-valued analog to discrete local context unit firing patterns is a uniform firing function of the form  $f_i(t) = \{1 \text{ if } \alpha_i \leq t \leq \beta_i, 0 \text{ otherwise}\}$ , with  $\alpha_i, \beta_i$ , and  $t$  real valued. Alternatively, a continuous-valued local context neuron could exhibit peak firing at a specific time point (or “center”) in the sequence with a more gradual decay in firing as the network state moves away from this center.

To study the difference between these two firing models (uniform and decaying) in the continuous case, we can specify the firing function by using clipped polynomial expressions of the form:

$$f_i(t) = \begin{cases} 1 - \left(\frac{t - \mu_i}{\sigma_i}\right)^{2k} & , \quad \mu_i - \sigma_i \leq t \leq \mu_i + \sigma_i \\ 0 & , \quad \text{otherwise} \end{cases} \quad (3.14)$$

While equation 3.14 is less familiar in the current context than more traditional gaussian firing models, it gives a more satisfying comparison between uniform and decaying firing models. Rather than using starting points and ending points to parameterize cell firing, each cell  $i$  is now characterized by a center,  $\mu_i$ , and a half-width scaling term,  $\sigma_i$ . Cells characterized by equation 3.14 reach peak firing rates at  $\mu_i$ ; firing decays over the length  $\sigma_i$  as the network moves away from the center, and the general shape of this decay is determined by (positive, integer-valued)  $k$ . In the limit as  $k$  approaches  $\infty$ , equation 3.14 becomes a uniform firing function with starting point  $\mu_i - \sigma_i$  and ending point  $\mu_i + \sigma_i$ . Thus, we can easily contrast uniform versus decaying firing models by comparing the synaptic weight distribution for low and high values of  $k$ . Figures 4a and 4b plots this firing function for various values of  $k$  and  $\sigma$ .

Picking up where we began in section 3.2.1, we can develop the synaptic weight distribution for the case of these continuous-valued firing functions using the same assumptions we used above (as enumerated in Table 1). Again, convergence assumptions (Assumption 2 in Table 1) imply that the synaptic weight  $w_{ij}$  takes on the correlation  $E[f_j(t)f_i(t-1)]/E[f_j(t)]$  (for a more precise justification in this continuous case, see Geman, 1979). For the sake of brevity, we assume that  $\sigma_i = \sigma$  is the same for each neuron (this is essentially like the simplification technique we use in section 3.2.3, where we assume that all context units have the same lifetime). As a result,

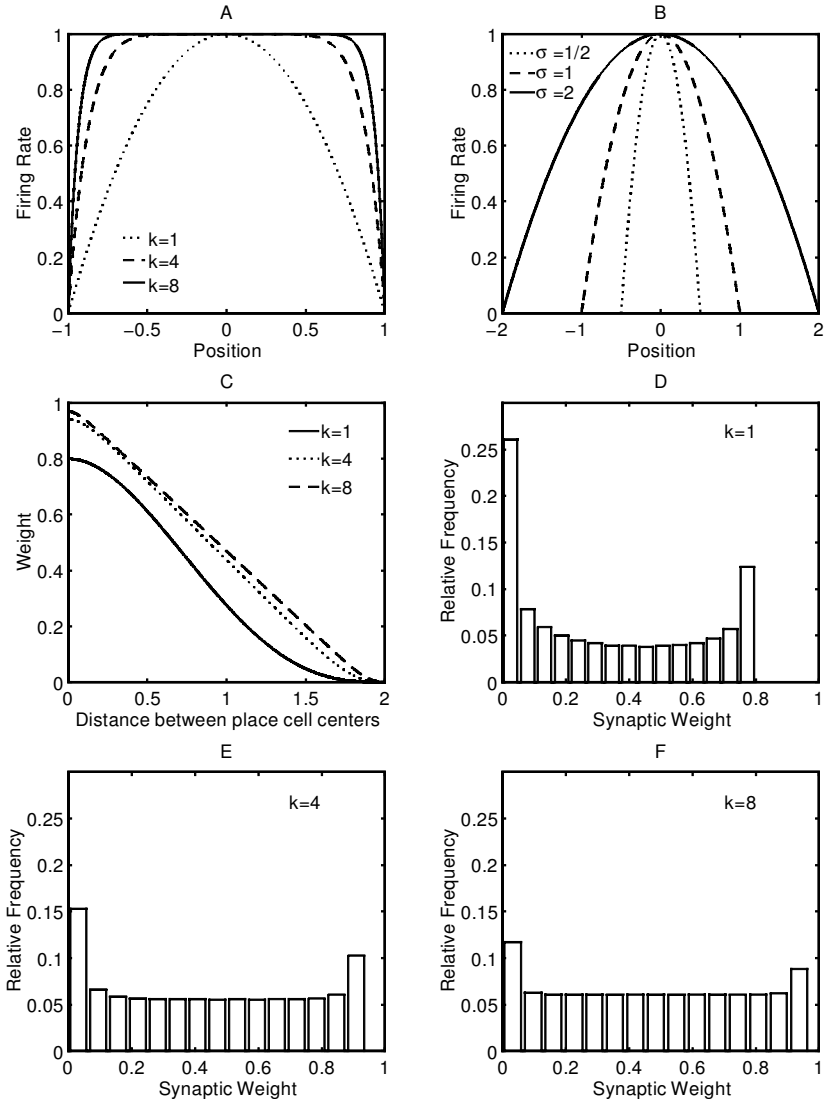
$$1/E[f_j(t)] = \frac{2k+1}{4k\sigma}$$

is a constant, which we call  $r$ . With  $\sigma$  fixed, then, we are interested in how the correlation  $rE[f_j(t)f_i(t-1)]$  varies as a function of the context unit centers,  $\mu_i$  and  $\mu_j$ . By applying Approximations 1 and 2, we assume that context unit centers are uniformly distributed across the sequence and note that, by virtue of independence,  $E[f_j(t)f_i(t-1)] = E[f_j(t)f_i(t)]$ . We denote the distance between place cell centers  $\mu_i$  and  $\mu_j$  as  $d_{ij} = |\mu_i - \mu_j|$ . By again ignoring end effects (Approximation 3), we obtain the synaptic weight as a function of the distance between place cell centers (the weight-distance function):

$$\begin{aligned} w_{ij} &= r \int_{d_{ij}-\sigma}^{\sigma} f(t)f(t-d_{ij}) dt \\ &= r \int_{d_{ij}-\sigma}^{\sigma} \left(1 - \left(\frac{t}{\sigma}\right)^{2k}\right) \left(1 - \left(\frac{t-d_{ij}}{\sigma}\right)^{2k}\right) dt, \end{aligned} \quad (3.15)$$

in the case that  $d_{ij} < 2\sigma$ ; otherwise,  $w_{ij} = 0$ . The advantage of using polynomials rather than gaussian functions becomes clear here because these integrals can easily be closed. Equation 3.15 can be rewritten as a closed function of  $\sigma$ ,  $d_{ij}$ , and  $k$  with the binomial theorem, though the resulting

series is a bit cumbersome for general  $k$ . Nevertheless, it is simple to show that the series generated by this expansion reduces to  $\sigma - \frac{1}{2}d_{ij}$  in the limit as  $k \uparrow \infty$ . That is, as the firing function approaches a uniform distribution, the synaptic weight between two cells becomes a linear function of the distance between their centers. On the other hand, for low values of  $k$ , the synaptic weight between two neurons is not a linear function of distance but is in-



stead a decaying, polynomial function of distance (e.g., for  $k = 1$ ,  $\sigma = 1$ , a fifth-degree polynomial:  $w_{ij} = -\frac{d_{ij}^5}{40} + \frac{d_{ij}^3}{2} - d_{ij}^2 + \frac{4}{5}$ ). Figure 4a and 4b plots the synaptic weight as a function of distance for  $\sigma = 1$  and various values of  $k$ .

The difference between these two classes of synaptic weight-distance curves (linear versus polynomial) is important because the synaptic weight distribution depends essentially on the rate of change of the synaptic weight function with respect to the distance between place cell centers. Before we take this last step, however, we point out that the linear version of this function (which arises from the uniform firing model) is essentially the result we have quantified for our McCulloch-Pitts hippocampal model, albeit in continuous form (this can be seen from equations 3.4 and 3.5, noting that  $\text{overlap}_{ij}$  is the discrete equivalent of  $K - d_{ij}$  with  $K$  some constant). This similarity leads to the reassuring conclusion that the shape of the synaptic weight distribution is consistent whether the neurons are continuous or binary.

Once the synaptic weight has been expressed as a function of the distance between place cell centers (the weight-distance function), all that remains is to apply the assumption that distance between place cell centers is uniformly distributed throughout the network (equivalent to Approximation 2 in a one-dimensional world). This will specify the synaptic weight distribution. Intuitively, it would appear that since this distance is uniformly distributed and the weight-distance function is monotonically decreasing, the density of the weight distribution will be relatively higher at those weights where the weight-distance function is decreasing slowly. Thus, it makes sense that a linear weight-distance function (with a constant rate of change) would give rise to a uniform weight distribution. Similarly, for the type of

---

Figure 4: *Facing page*. Weight analysis for the polynomial firing rate function, equation 3.14,  $f_i(t) = 1 - (\frac{t - \mu_i}{\sigma})^{2k}$ . Here  $\mu_i$  is 0. (A)  $f_i(t)$  for various values of  $k$ . The parameter  $\sigma$  is held constant at 1. Notice how the  $f_i(t)$  approaches the uniform firing function as  $k$  increases. (B)  $f_i(t)$  for various values of  $\sigma$ ;  $k$  is held constant at 1. (C) The synaptic weight as a function of the distance between place centers for two cells (the weight-distance function for two neurons), as in equation 3.15. The weight-distance function is plotted for several values of  $k$ , with  $\sigma = 1$ . Notice how the function becomes linear as  $k$  increases. (D–F) Nonzero synaptic weight histograms were generated numerically for  $k$  of 1, 4, and 8. A uniform distribution of the distance between place cell centers (as would be the case if place cells were uniformly distributed in a one-dimensional world), and no end effects, were assumed. A set of distances evenly spaced (between 0 and 2) was transformed by the appropriate weight-distance function, and the histograms were constructed from this transformed set. Notice how the nonzero synaptic weight distribution becomes increasingly uniform as  $k$  increases.

weight-distance function that arises for a gradually decaying firing function such as  $k = 1$  (and also for a gaussian firing function; see below), the weight distribution would be relatively higher for high and low weights than for intermediate weights, giving rise to a U-shaped type of distribution.

More formally, we partition the zero and nonzero synaptic weight distributions. Generalizing, suppose we have a weight-distance function,  $g$ , such that  $w_{ij} = g(d_{ij})$ , which is strictly decreasing (hence invertible) for  $d_{ij} \leq 2\sigma$  and zero valued for  $d_{ij} > 2\sigma$ . Distances are distributed uniformly over half the sequence length (density  $2/S$ ). In the first case, we have  $P(w_{ij} = 0) = P(d_{ij} > 2\sigma) = 1 - 4\sigma/S$ , where  $S$  is again the sequence length. For the nonzero weights,

$$\begin{aligned} P(0 < w_{ij} \leq t) &= P(0 < g(d_{ij}) \leq t) = P(g^{-1}(t) \leq d_{ij} < g^{-1}(0)) \\ &= \int_{g^{-1}(t)}^{g^{-1}(0)} 2/S dx = -2/S \int_0^t \frac{dg^{-1}(x)}{dx} dx. \end{aligned} \quad (3.16)$$

The density function of the nonzero synaptic weight distribution is thus

$$\frac{-2}{S} \frac{dg^{-1}(x)}{dx},$$

proportional to the first derivative of the inverse of the weight-distance function.

To determine the synaptic weight distribution for other cell firing models, all that is needed is an expression of the synaptic weight as a function of the distance between place cell centers, such as is computed in equation 3.15. Provided this function is decreasing (the only imaginable case in pure place identification), the density of the synaptic weight distribution will be proportional to the first derivative of the inverse of this function; geometrically, this means that the density will be relatively higher at those places where the weight-distance function is changing relatively slowly, so it should be possible to draw qualitative conclusions from the weight-distance function itself.

Even for the simplest case,  $k = 1$ ,  $\sigma = 1$ , the weight-distance function arising from equation 3.14 is a fifth-degree polynomial (above), which we did not invert. However, the weight-distance function (see Figure 4C) itself indicates that its rate of change is smaller for low and high distances (by implication, high and low weights) than for intermediate distances (weights). This result suggests a U-shaped distribution for the nonzero synaptic weights, as is confirmed by numerical calculations (see Figure 4D). The difference between trough and peak, however, is not large, so the uniform distribution remains a good approximation. As  $k$  increases (i.e., as the firing function becomes more uniform), however, the weight-distance function becomes more linear (approaching  $\sigma - \frac{1}{2}d_{ij}$ ), and gives rise to a uniform nonzero distribution. This can be demonstrated analytically, in the

limit as  $k \uparrow \infty$ . Figure 4 illustrates these results from numerical calculations for various values of  $k$ .

In summary, a uniform place-identification firing model gives rise to a uniform nonzero synaptic weight distribution in the case of continuous neurons (and continuous time), as well as in the discrete case of the McCulloch-Pitts model. If a decaying firing model is used, the synaptic distribution rises at the high and low ends of the weight spectrum, becoming U shaped; however, even in this case, the uniform distribution remains a good approximation. The result for the zero weights remains essentially the same.

*3.4.2 Gaussian place cell firing.* Extending the previous analysis to the case of gaussian place cell firing is straightforward. Suppose now the firing function for cell  $i$  is a gaussian function,

$$f_i(t) = e^{-\frac{(t-\mu_i)^2}{2\sigma^2}}, \quad (3.17)$$

with center  $\mu_i$ . For the sake of simplicity, we again assume that the variance  $\sigma^2$  is the same for all cells. We develop the weight-distance function in the same way, with  $d_{ij}$  again the distance between place cell centers:

$$w_{ij} = \frac{\int f_i(x)f_j(x) dx}{\int f_j(x) dx} = \frac{\int f(x)f(x-d_{ij}) dx}{\int f_j(x) dx}. \quad (3.18)$$

The numerator of equation 3.18 reduces with equation 3.17,

$$\begin{aligned} \int f(x)f(x-d_{ij}) dx &= \int \exp\left[\frac{1}{2\sigma^2}(-x^2 - (x-d_{ij})^2)\right] dx \\ &= \exp\left(-\frac{d_{ij}^2}{4\sigma^2}\right) \int \exp\left[-\frac{(x-\frac{1}{2}d_{ij})^2}{\sigma^2}\right] dx, \end{aligned} \quad (3.19)$$

which gives

$$w_{ij} = \frac{\exp\left(-\frac{d_{ij}^2}{4\sigma^2}\right) \int \exp\left[-\frac{(x-\frac{1}{2}d_{ij})^2}{\sigma^2}\right] dx}{\int \exp\left[-\frac{x^2}{2\sigma^2}\right] dx} \approx \frac{\sqrt{2}}{2} \exp\left(-\frac{d_{ij}^2}{4\sigma^2}\right). \quad (3.20)$$

The natural course would be to take these integrals over the sequence length. But if instead we approximate with integrals extending to infinity, the error functions disappear, and the weight distribution becomes analytically tractable (this approximation follows from Approximation 3, see section 3.3.5) because the ratio of the two integrals then becomes  $\sqrt{2}/2$ . The weight-distance curve is thus itself a gaussian function.

Following equation 3.16, we can invert and differentiate equation 3.20, to give a weight density function

$$f_W(x) = \frac{2}{S} \frac{\sigma}{x\sqrt{-\ln(\sqrt{2}x)}}, \frac{\sqrt{2}}{2} \exp\left(-\frac{S^2}{16}\right) < x < \frac{\sqrt{2}}{2}, \quad (3.21)$$

which is (as we would by now expect) U shaped (see Figure 5).

**3.4.3 Stochastic binary neurons.** Returning to the case of discrete time and binary firing, what if the firing functions are stochastic rather than deterministic? If we have firing functions of the form  $P(z_i(t) = 1) = f_i(t)$ , then the situation is the same as that which develops out of equation 3.2. The analysis is effectively the same as for the continuous cases, though the calculations now involve finite sums and a discrete space of possibilities. The case of uniform firing (i.e., local context units) is covered by equation 3.8 and gives rise, via approximation, to a uniform weight distribution (see equations 14a–c). For alternative stochastic firing distributions (such as mean peaked, decaying), we could approximate the firing function with a continuous function and use the techniques of the previous two sections to derive a continuous weight distribution that approximates the discrete one. Thus, we would again expect a U-shaped weight distribution for a mean-peaked firing function such as a discretized gaussian.

**3.4.4 A two-dimensional world.** The assumption that distance between place cell centers is uniformly distributed only follows from the assumption that place cells are uniformly distributed (e.g., Approximation 2) in a one-dimensional world. In a two-dimensional world, however, the number of place fields farther from a central locus is greater than the number of place cells closer. Briefly, suppose that place cells are uniformly distributed on a planar circle with radius  $R$ , and consider a particular place cell at the center of the circle. We want to specify the distribution of the distances between the place cell at the center of the circle and other place cells in the circle (we use the center place cell here and assume that this distribution would be the same for all other place cells—this is strictly true only if there are no end effects). Referring to the distance variable as  $d$ , we have  $P(d \leq t) = \pi t^2 / \pi R^2 = t^2 / R^2$ , which implies a density function  $f_d(t) = 2t / R^2$ . The density of the distribution actually increases in direct proportion with the distance.

This modification alters the derivation of weight density given by equation 3.16. Instead,

$$\begin{aligned} P(0 < w_{ij} \leq t) &= P(0 < g(d_{ij}) \leq t) = P(g^{-1}(t) \leq d_{ij} < g^{-1}(0)) \\ &= \int_{g^{-1}(t)}^{g^{-1}(0)} \frac{2x}{R^2} dx = \frac{-2}{R^2} \int_0^t g^{-1}(u) \frac{dg^{-1}(u)}{du} du, \end{aligned} \quad (3.22)$$

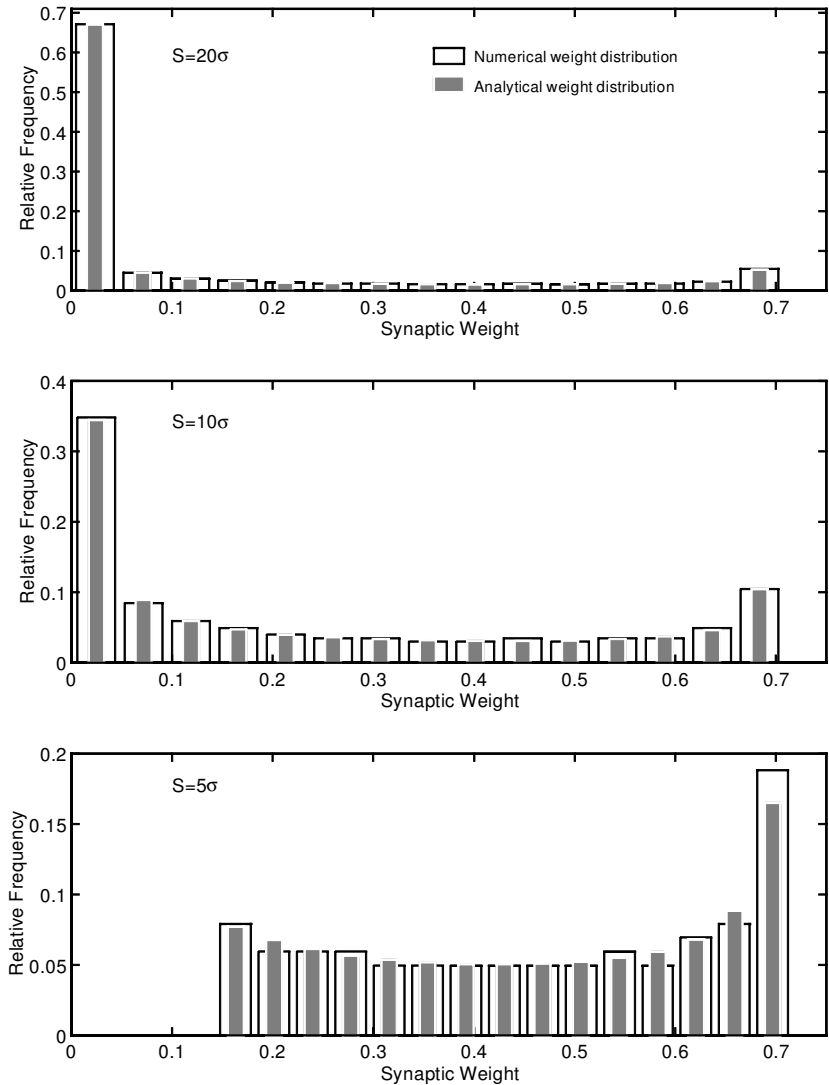


Figure 5: Weight analysis for the gaussian firing rate function, equation 3.17. The numerical weight distributions were computed by constructing a set of distances evenly spaced between 0 and  $S/2$ , where  $S$  was the sequence length. This set was transformed by the weight-distance function generated by the middle equation in equation 3.20; the integrals were computed numerically. The numerical histograms were constructed from this transformed set. The analytical histograms were constructed by integrating the weight density function specified by equation 3.21 over the bin widths. Comparisons were generated separately for sequence length ( $S$ ) values of  $5\sigma$ ,  $10\sigma$ , and  $20\sigma$ . Note that the scale on the abscissa differs for each histogram.

using the substitution  $u = g(x)$ , where  $g$  is a strictly decreasing weight-distance function (again, if  $g$  is not strictly decreasing everywhere, the weight distribution would need to be put together from partitions).

The weight density function in this two-dimensional case is thus  $f_w(t) = \frac{-2}{R^2} g^{-1}(x) \frac{dg^{-1}}{dx}$ , which differs from the density in the one-dimensional case, equation 3.16, by the factor  $g^{-1}(x)$ . This difference reflects the factor that more place cells are actually at farther distances than closer ones if place cells are uniformly distributed in a two-dimensional world.

Extending this analysis to the cognitive mapping problem requires specifying the firing function of a place cell in two dimensions, deriving the weight-distance function via the integration techniques described in the previous two sections, and then using equation 3.22. Due to the great diversity of cell firing functions and weight-distance functions that have been measured or hypothesized for hippocampal place cells (Blum & Abbot, 1996; O'Keefe & Burgess, 1996; Burgess & O'Keefe, 1996; Touretzky, Redish, & Wan, 1993; Tsodyks & Sejnowski, 1995; Muller, Kubie, & Saypoff, 1991; Muller & Stead, 1996), an explicit analysis of the various cases falls outside the scope of this article; still, something in general can be said.

This two-dimensional effect can be simply though crudely exemplified by a linear weight-distance function (a firing function uniform on a circle gives rise to a weight-distance function that is very nearly linear). Assume the weight-distance function  $g(d_{ij}) = M - d_{ij}$  if  $d_{ij} \leq M$ , and is 0 otherwise. Then the nonzero synaptic weight density, given by equation 3.22,  $f_w(x) = \frac{2}{R^2}(M - x)$ , is a linear, decreasing distribution—in contrast to the uniform distribution that arises in a one-dimensional world.

**3.4.5 Summary.** The purpose of this section has been to demonstrate that the computational techniques used to quantify the synaptic weight distribution in our hippocampal model generalize, using analogous assumptions and approximations, to cell firing models other than those of deterministic, binary, discrete local context units. While it is useful to understand how these computations generalize, it bears reiterating that this does not imply that the model generalizes. Whether a model can be constructed that evolves a coding similar to any of the cases we have discussed and is well characterized by the assumptions and approximations we use (particularly Approximation 1, that coding is independent of connectivity) is an issue completely divorced from these generalizations.

For all the cases analyzed in detail, the essential result of a nongaussian synaptic weight distribution remains. While these distributions were not always uniform, the uniform distribution is a good approximation.

## 4 Discussion

---

The purpose of this analysis was to understand how the relationship between the input environment and the distribution of synaptic weights is

affected by two important phenomena: Hebbian synaptic modification and the emergence of the local context unit-based coding scheme.

We find that this relationship imposes a distinctive constraint on the distribution of the synaptic weights under the sufficient condition that context units form. The actual distributions are, to a remarkably large extent, independent of the details of input firing patterns. Using equation 14a–c, these distributions can instead be predicted as a function of the average context unit lifetime (which depends on more general characteristics of the input environment) with minimal inaccuracy. Even more simply, by using the relationship between average context unit lifetime,  $E[\ell]$ , capacity, and activity described by equation 2.3, it is possible to approximate these distributions with even more rudimentary knowledge. That is, if local context codes form and if average activity is known or preset, the distribution is largely independent of the environment, according to equation 3.12.

The surprise of these results suggests a contradiction to two assumptions underlying our original intuitions concerning the dynamical properties of this network: (1) an expectation of a typical gaussian-type weight distribution arising from a first-approximation assumption of independent neuronal firing, and (2) the general idea that Hebbian weight modification imposes the statistical characteristics of the environment on the synapses.

Our intuitive expectation of a gaussian weight distribution builds on a supposition of independent neuronal firing. As a simplified example, suppose that each neuron fires independently  $Sa$  times, so that the average network activity is  $a$ . Again, convergence assumptions imply that  $w_{ij}$  will approach  $P(z_i(t-1) | z_j(t))$ . If  $k_{ij}$  is the number of time steps of synchronous firing (like overlap $_{ij}$ ), it follows that

$$P(k_{ij} = m) = \frac{\binom{Sa}{m} \binom{S-Sa}{Sa-m}}{\binom{S}{Sa}}; \quad (4.1)$$

that is,  $k_{ij}$  has a hypergeometric distribution. But since  $w_{ij} = k_{ij}/Sa$ , this implies  $w_{ij}$  is also hypergeometrically distributed with mean  $a$  and variance  $(1-a)^2/(S-1)$ . As  $S \rightarrow \infty$ , the central limit theorem implies that the standardized distribution becomes gaussian, not uniform, as in our results. Moreover, the distribution converges to an impulse function at the mean,  $a$ , as  $S \rightarrow \infty$  as well, as implied by the variance term. Thus  $P(w_{ij} = 0)$  converges to 0, not to the ca.  $1-2a$  of our results. The weight distribution in this formerly intuitive case is therefore quite different from the distribution shown to exist when local context units characterize cell firing.

The difference between the two coding schemes that produce these very different distributions stems from a difference in the input environments and the effect of synaptic modification. The case of independent activity patterns is characteristic of a structureless, informationless environment, while local context unit-type activity arises (in the model) from predictable correlations in the sequential repetition of the input.

Understanding the basis for our earlier, false intuition leads us directly to the issue of biological realism. The world and the brain is some hybrid combination of the two extremes—the one of our initial intuition and the one explained previously. Although the relative contribution of each of these two input types is an open question, it will always be the challenge to the information processor to extract whatever signal there is from amid however much noise there is. So it would make sense that the relationship between the synaptic weight structure and coding scheme of a more natural sequence-learning and predicting network, such as is found in the mammalian brain, lies somewhere between the extremes of purely deterministic and uniform, local context unit-based codings and the purely noisy and gaussian, independent firing patterns. Thus, we suggest that the next step in the synaptic weight theory presented here requires relaxing some, *but not all*, of its deterministic rigor.

The fact that the input environments simulated here are less noisy than in more realistic, biological cases makes it all the more surprising that the synaptic weight distribution turns out to be independent of the input environment, on the condition that context units form. As a result, we suspect that some of the more qualitative results of the deterministic case analyzed here are meaningful.

Among these results, the large number of synapses that are driven to zero weight is hypothesized to be a general characteristic of these networks. Consequently, the theoretical link between zero weights and low activity developed as equation 3.12,  $P(w_{ij} = 0) \sim 1 - 2a$ , could provide a unification of two well-known and complementary characteristics of forebrain cortical systems: low connectivity and low activity.

Generalizing the uniform distribution of nonzero synaptic weights is a more complicated question, and this complication formed the basis for the extensions presented in section 3.4. Here, we demonstrate how the results can extend to the cases of stochastic and continuous firing neurons, using only assumptions and approximations equivalent to those already introduced (see Table 1). The essential result of a nongaussian weight distribution still obtains, but with an additional qualification: the resulting nonzero weight distribution may be either uniform or U shaped, depending on the “place cell” firing function assumed. In both cases, deriving the weight distribution from the firing function, for both analytical and numerical analysis, depends on the firing function in the same straightforward way. Moreover, in the cases we studied, the uniform distribution of weights remained a good approximation even for the distributions predicted to be U shaped.

This leads directly to the question of whether the assumptions outlined in Table 1 are biologically reasonable. Section 3.3 discusses the issues surrounding this question. As this discussion indicates, our position is that all of these approximations are computationally reasonable, with the exception of Approximation 1 (see section 3.3.3), which states, in effect, that a

neuron's role in the coding is independent of its connectivity. The success of this particular approximation in the computational simulations is actually a bit surprising. In this context, an important result of our model is that a self-organizing recurrent model of the CA3 region can develop local context units for which neuronal correlations (i.e., the relationship between two connected neurons which depends on the respective  $\alpha$ 's and  $\beta$ 's) can be very effectively approximated by an assumption of independence to the variable of connectivity. This result is important because context unit place cell codings are not imposed on our model, but rather develop out of a self-organizing process. Therefore, it is surprising that the model's principal source of randomness, the connectivity matrix, can be assumed to have no influence on this process when the weight distribution is computed.

There are several interesting cases, in addition, in which this coding-connectivity independence assumption is necessarily valid and in which we would expect our extended results to hold. First are the models that implicitly presuppose independence between coding and connectivity, by virtue of either full connectivity (e.g., Blum & Abbott, 1996) or externally originated place cells imposed, independently, on a randomly connected network (e.g., Muller & Stead, 1996).

Second, and perhaps of some interest to neurophysiologists, the distributions of weights could just as well have been the distribution of pairwise cell firing,  $E[z_i z_j]/E[z_j]$  (and this is true regardless of whether the coding-connectivity independence assumption is valid). The motivation to measure these correlations has begun to be shared by experimentalists, and the distributions published in the recent experimental paper of Hampson, Byrd, Konstantopoulos, Bunn, & Deadwyler (1996; see, e.g., their Figures 3C and 3D) are very similar to the sorts of distributions described here.<sup>1</sup>

Finally, we note that the description of synfire chains as models of cortical function suggested by Abeles (1991), and similar lines of thought arising from other researchers, has recently inspired a great deal of research on sequence-producing recurrent neural networks. The firing patterns in these cortically inspired networks are characterized by subpopulation neuronal groups (synfire links) activating in sequential fashion (Abeles 1991; Bienenstock, 1995; Hertz & Prügel-Bennet, 1996; Herrmann, Hertz, & Prügel-Bennet, 1995; Griniasty et al., 1993; Amit et al., 1994; Brunel, 1996; Chover, 1996). Thus we believe our results are not limited to hippocampal models. Rather they may be applied to a variety of models currently being studied in the computational literature.

---

<sup>1</sup> Hampson et al. refer to these as *cross-correlations*. They measure the distribution of  $E[z_i z_j]$  rather than the  $E[z_i z_j]/E[z_j]$  distribution we calculate. The difference in the distributions,  $1/E[z_j]$ , is essentially one of scale (our measure is scaled for the average activity of the postsynaptic neuron), not of shape.

## Appendix

---

The purpose of this appendix is to make explicit the calculations that produce  $P(w_{ij} = x | \ell_i, \ell_j)$ . To do this calculation, the  $(\ell_i, \ell_j)$  space was partitioned into the three cases:  $\ell_i = \ell_j$ ,  $\ell_i > \ell_j$ , and  $\ell_i < \ell_j$ . Within these partitions, we subpartition the cases as a function of starting point and lifetime, in order to simplify the min and max functions in the  $\text{overlap}_{ij}$  definition; this is made easier if we consider the relationship separately depending on which context unit (the pre- or the post-) starts firing first. Finally, these subpartitions are reduced to joint probabilities and then unified to remove the subpartitions. Among these three cases, the conditional weight distribution is structurally very similar, and so we are able to unify the three cases into a single function. This is the result we present as equation 3.8.

**A.1 Case I:**  $\ell_i = \ell_j = \ell$ . In this case equation 3.7 reduces to:

$$\text{overlap}_{ij} = \max\{\min\{\alpha_i + \ell + 1, \alpha_j + \ell\} - \max\{\alpha_i + 1, \alpha_j\}, 0\}, \quad (\text{A.1})$$

in which case we can further partition the possibilities into Case Ia, when the presynaptic context unit starts firing before the postsynaptic, and vice versa in Case Ib.

**A.2 Case Ia:**  $\alpha_j \geq \alpha_i + 1$ . So that  $\max\{\alpha_i + 1, \alpha_j\} = \alpha_j$  and  $\min\{\alpha_i + \ell + 1, \alpha_j + \ell\} = \alpha_i + \ell + 1$ . This gives  $\text{overlap}_{ij} = \max\{\alpha_i + \ell + 1 - \alpha_j, 0\}$  and allows us to conclude:

$$P(\text{overlap}_{ij} = 0, \alpha_j \geq \alpha_i + 1) = P(\alpha_i + \ell + 1 - \alpha_j \leq 0, \alpha_j \geq \alpha_i + 1) \quad (\text{A.2a})$$

$$P(\text{overlap}_{ij} = \ell, \alpha_j \geq \alpha_i + 1) = P(\alpha_j = \alpha_i + 1, \alpha_j \geq \alpha_i + 1) \quad (\text{A.2b})$$

$$P(\text{overlap}_{ij} = x, \alpha_j \geq \alpha_i + 1) = P(\alpha_i + \ell + 1 - \alpha_j = x, \alpha_j \geq \alpha_i + 1), \\ 0 < x < \ell \quad (\text{A.2c})$$

for the cases of no overlap, complete overlap, and partial overlap, respectively.

**A.3 Case Ib:**  $\alpha_j < \alpha_i + 1$ . Then  $\max\{\alpha_i + 1, \alpha_j\} = \alpha_i + 1$  and  $\min\{\alpha_i + \ell + 1, \alpha_j + \ell\} = \alpha_j + \ell$ . So  $\text{overlap}_{ij} = \max\{\alpha_j + \ell - \alpha_i + 1, 0\}$ , and, once again, we have

$$P(\text{overlap}_{ij} = 0, \alpha_j < \alpha_i + 1) = P(\alpha_j + \ell - \alpha_i + 1 \leq 0, \alpha_j < \alpha_i + 1) \quad (\text{A.3a})$$

$$P(\text{overlap}_{ij} = \ell, \alpha_j < \alpha_i + 1) = P(\alpha_j = \alpha_i + 1, \alpha_j < \alpha_i + 1) = 0 \quad (\text{A.3b})$$

$$P(\text{overlap}_{ij} = x, \alpha_j < \alpha_i + 1) = P(\alpha_i + \ell + 1 - \alpha_j = x, \alpha_j < \alpha_i + 1),$$

$$0 < x < \ell \tag{A.3c}$$

for the cases of no overlap, complete overlap, and partial overlap, respectively.

Since Cases Ia and Ib partition the event space, we can generate the  $\text{overlap}_{ij}$  distribution by combining equations A.2a-c and A.3a-c. The calculation of  $P(\text{overlap}_{ij} = 0)$  is presented as an example:

$$P(\text{overlap}_{ij} = 0) = P(\alpha_j \leq \alpha_i - \ell + 1, \alpha_j < \alpha_i + 1)$$

$$+ P(\alpha_j \geq \ell + 1 + \alpha_i, \alpha_j \geq \alpha_i + 1)$$

$$= P(\alpha_j \leq \alpha_i - \ell + 1) + P(\alpha_j \geq \ell + 1 + \alpha_i)$$

$$= 1 - P(\alpha_i - \ell + 1 < \alpha_j < \alpha_i + \ell + 1). \tag{A.4}$$

Equation A.4, as well as its analog for the other cases, can be computed by assuming independence of  $\alpha_i$  and  $\alpha_j$  and invoking the assumption of uniform distribution of  $\alpha$ 's (so that, explicitly,  $P(\alpha_i = t) = 1/(S - \ell_i + 1)$  if  $1 \leq t \leq S - \ell_i + 1$ ):

$$P(\text{overlap}_{ij} = 0)$$

$$= 1 - P(\alpha_i - \ell + 1 < \alpha_j < \alpha_i + \ell + 1),$$

[which we break down by again partitioning into conditional relationships]

$$= 1$$

$$- \sum_{x=1}^{S-\ell+1} \sum_{y=1}^{S-\ell+1} P(\alpha_i - \ell + 1 < \alpha_j < \alpha_i + \ell + 1 \mid \alpha_i = x, \alpha_j = y)$$

$$\times P(\alpha_i = x, \alpha_j = y)$$

[Since  $x$  and  $y$  are constants (in time-step units),  $P(\alpha_i - \ell + 1 < \alpha_j < \alpha_i + \ell + 1 \mid \alpha_i = x, \alpha_j = y)$  is an indicator function:  $P(\alpha_i - \ell + 1 < \alpha_j < \alpha_i + \ell + 1 \mid \alpha_i = x, \alpha_j = y) = 1$  if and only if  $x - \ell + 1 < y < x + \ell + 1$ , provided that  $P(\alpha_i = x) > 0$  and  $P(\alpha_j = y) > 0$ . Otherwise, it takes value 0. We rewrite using the notation  $I(A)$  as the indicator function of event  $A$  and apply the independence assumption for  $\alpha_i$  and  $\alpha_j$ .]

$$= 1 - \sum_{x=1}^{S-\ell+1} \sum_{y=1}^{S-\ell+1} I(x - \ell + 1 < y < x + \ell + 1) P(\alpha_i = x) P(\alpha_j = y)$$

$$= 1 - \sum_{x=1}^{S-\ell+1} \sum_{y=1}^{S-\ell+1} I(x - \ell + 1 < y < x + \ell + 1) \frac{1}{S - \ell + 1} \frac{1}{S - \ell + 1}$$

[The calculation of  $\sum_y I(x - \ell + 1 < y < x + \ell + 1)$  here is complicated by the end-effect situation. In the ideal case, the event  $\{x - \ell + 1 < y < x + \ell + 1\}$  would be measurable without respect to the particular value that  $x$  takes, as would appear from the event itself. However, the probability distribution of  $\alpha = y$  is additionally constrained by the condition  $1 \leq \alpha \leq S$ , and this complicates the calculation. We avoided complicating the theory by ignoring such a possibility, thus assuming  $\ell_i < x < S - \ell_i$ , or, effectively, that the presynaptic context unit was not near the beginning or the end of the sequence. In light of this assumption, noting that  $x$  is a constant within the summation,  $\sum_y I(x - \ell + 1 < y < x + \ell + 1)$  amounts to a count of the number of values of  $y$  satisfying  $x - \ell + 1 < y < x + \ell + 1$  for any constant  $x$ , but this is just  $2\ell - 1$ .]

$$\begin{aligned}
 &= 1 - \left( \frac{1}{S - \ell + 1} \right)^2 \sum_{x=1}^{S-\ell+1} 2\ell - 1 \\
 &= 1 - \frac{2\ell - 1}{S - \ell + 1} \\
 &= \frac{S - 3\ell + 2}{S - \ell + 1} \tag{A.5}
 \end{aligned}$$

The procedure employed to calculate  $P(\text{overlap}_{ij} = 0 \mid \ell)$  above applies generically for the remaining cases  $0 < \text{overlap}_{ij} \leq \ell$ .

Given this conditional probability, it is a simple matter to compute  $P(w_{ij} \mid \ell)$  with equation 3.6.

**A.4 Case II:**  $\ell_i > \ell_j$ . The procedure for the remaining cases is generically similar to Case I, so here we only point out the differences and show the results of the calculations. The most significant characteristic distinguishing Cases II and III from Case I is that the information contained in  $\ell_i$  and  $\ell_j$  cannot be simplified to a single variable; a conditional probability can only be generated for both  $\ell_i$  and  $\ell_j$ . This affects the indices of summation in the analogous calculation producing equation A.5 as well as the defining terms of the end-effect error. However, in the former case, the indices are cancelled out by the  $(\alpha_i, \alpha_j)$  independence assumption, and, in the latter, the end error can be assumed away as before. The  $\text{overlap}_{ij}$  can be calculated from  $(\alpha_i, \beta_i)$  and  $(\alpha_j, \beta_j)$  as above, and the same computations applied:

$$P\left(w_{ij} = \frac{x}{\ell_j} \mid \ell_i, \ell_j\right) = \begin{cases} \frac{S-2\ell_i-\ell_j+2}{S-\ell_j+1} & , \quad x = 0 \\ \frac{2}{S-\ell_j+1} & , \quad 0 < x < \ell_j \\ \frac{\ell_i-\ell_j+1}{S-\ell_j+1} & , \quad x = \ell_j \end{cases} \tag{A.6}$$

**A.5 Case III:**  $\ell_i < \ell_j$ . The same reasoning applies if  $\ell_i < \ell_j$ . This is guaranteed by the symmetry of the  $\ell_i < \ell_j$  and  $\ell_i > \ell_j$  cases if, as we have assumed, the two variables are independent. In this case, however, since  $0 < \text{overlap}_{ij} < \ell_i$ , it follows  $0 < w_{ij} < \ell_i/\ell_j$ , and thus  $P(w_{ij} = 1) = 0$  necessarily:

$$P\left(w_{ij} = \frac{x}{\ell_j} \mid \ell_i, \ell_j\right) = \begin{cases} \frac{s-2\ell_i-\ell_i+2}{s-\ell_j+1}, & x = 0 \\ \frac{2}{s-\ell_j+1}, & 0 < x < \ell_i \\ \frac{\ell_i-\ell_i+1}{s-\ell_j+1}, & x = \ell_i \end{cases} \quad (\text{A.7})$$

## Acknowledgments

---

This work was supported by NIH MH48161 and MH00622, EPRI RP8030-08 and Pittsburgh Supercomputing Center Grant BNS950001P to WBL, and by the Department of Neurosurgery, John A. Jane, chairman. We thank Xi-angbao Wu for his patient advice and assistance with the simulations and Nancy L. Desmond for her assistance with the manuscript.

## References

---

- Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral circuits*. Cambridge: Cambridge University Press.
- Amit, D. J., Brunel, N., & Tsodyks, M. V. (1994). Correlations of cortical Hebbian reverberations: Theory versus experiment. *Journal of Neuroscience*, *14*(11), 6435–6445.
- Bienenstock, E. (1995). A model of neocortex. *Network*, *6*(2), 179–224.
- Blum, K. I., & Abbott, L. F. (1996). A model of spatial map formation in the hippocampus of the rat. *Neural Computation*, *8*(1), 85–93.
- Brunel, N. (1996). Hebbian learning of context in recurrent neural networks. *Neural Computation*, *8*(8), 1677–1710.
- Burgess, N., & O'Keefe, J. (1996). Neuronal computations underlying the firing of place cells and their role in navigation. *Hippocampus*, *6*(6), 749–762.
- Chover, J. (1996). *Sequential recall*. Unpublished manuscript.
- Geman, S. (1979). Some averaging and stability results for random differential equations. *SIAM J. Appl. Math.*, *36*(1), 86–105.
- Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.
- Griniasty, M., Tsodyks, M. V., & Amit, D. J. (1993). Conversion of temporal correlations between stimuli to spatial correlations between attractors. *Neural Computation*, *5*(1), 1–17.
- Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*. Boston: D. Reidel.

- Hampson, R. E., Byrd, D. R., Konstantopoulos, J. K., Bunn, T., & Deadwyler, S. A. (1996). Hippocampal place fields: Relationship between degree of field overlap and cross-correlations within ensembles of hippocampal neurons. *Hippocampus*, 6, 281–293.
- Herrmann, M., Hertz, J., & Prügel-Bennet, A. (1995). Analysis of synfire chains. *Network*, 6, 403–414.
- Hertz, J., & Prügel-Bennet, A. (1996). Learning short synfire chains by self-organization. *Network*, 7(2), 357–363.
- Hirsh, R. (1974). The hippocampus and contextual retrieval from memory. *Behav. Biol.*, 12, 421–444.
- Kesner, R. P., & Hardy, J. D. (1983). Long-term memory for contextual attributes: Dissociation of amygdala and hippocampus. *Behav. Brain Res.*, 8, 139–149.
- Levy, W. B. (1982). Associative encoding at synapses. In *Proc. Fourth Ann. Conf. Cogn. Sci. Soc.* (pp. 135–136).
- Levy, W. B. (1989). A computational approach to hippocampal function. In R. D. Hawkins & G. H. Bower (Eds.), *Computational modeling of learning in simple neural systems* (pp. 243–305). Orlando, FL: Academic Press.
- Levy, W. B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus*, 6(6), 579–590.
- Levy, W. B., & Steward, O. (1979). Synapses as associative memory elements in the hippocampal formation. *Brain Res.*, 175, 233–245.
- Levy, W. B., & Steward, O. (1983). Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience*, 8(4), 791–797.
- Levy, W. B., & Wu, X. B. (1996). The relationship of local context codes to sequence length memory capacity. *Network*, 7(2), 371–382.
- Levy, W. B., & Wu, X. B. (1997). A simple, biologically motivated neural network solves the transitive inference problem. In *Proceedings of the 1997 International Conference on Neural Networks*, Vol. 1 (pp. 368–371).
- Levy, W. B., Wu, X. B., & Baxter, R. A. (1995). Unification of hippocampal function via computational/encoding considerations. In *Proceedings of the Third Workshop: From Biology to High-Energy Physics. Intl. J. Neural Sys.*, Vol. 6 (Suppl.) (pp. 71–80).
- Levy, W. B., Wu, X. B. & Tyrcha, J. M. (1996). Solving the transverse patterning problem by learning context present: A special role for input codes. In *INNS World Congress on Neural Networks* (pp. 1305–1309). Hillsdale, NJ: Erlbaum.
- McClelland, J., & Goddard, N. (1996). Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. *Hippocampus*, 6(6), 654–665.
- Minai, A. A., Barrows, G. L., & Levy, W. B. (1994). Disambiguation of pattern sequences with recurrent networks. In *INNS World Congress on Neural Networks* (Vol. 4, pp. 176–181). Hillside, NJ: Erlbaum.
- Muller, R. U., Kubie, J. L., & Saypoff, R. (1991). The hippocampus as a cognitive graph (abridged version). *Hippocampus*, 1(3), 243–246.
- Muller, R. U., & Stead, M. (1996). Hippocampal place cells connected by Hebbian synapses can solve spatial problems. *Hippocampus*, 6(6), 709–719.

- Nadel, L., & Willner, J. (1980). Context and conditioning: A place for space. *Physiol. Psych.*, *8*(2), 218–228.
- O'Keefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, *381*, 425–428.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- Thompson, L. T., & Best, P. J. (1989). Place cells and silent cells in the hippocampus of freely-behaving rats. *J. Neurosci.*, *9*(7), 2382–2390.
- Touretzky, D. S., Redish, A. D., & Wan, H. S. (1993). Neural representation of space using sinusoidal arrays. *Neural Computation*, *5*(6), 869–884.
- Treves, A., & Rolls, E. T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus*, *2*, 189–199.
- Treves, A., & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, *4*(3), 374–391.
- Tsodyks, M., & Sejnowski, T. (1995). Associative memory and hippocampal place cells. In *Proceedings of the Third Workshop: From Biology to High-Energy Physics. Intl. J. Network Sys.*, Vol. 6 (Suppl.) (pp. 81–86).
- Voydovic, J. T. (1996). Cell death in cortical development: How much? why? so what? *Neuron*, *16*, 693–696.
- Wu, X. B., Baxter, R. A., & Levy, W. B. (1996a). Context codes and the effect of noisy learning on a simplified hippocampal CA3 model. *Biol. Cybern.*, *74*, 159–165.
- Wu, X. B., Tyrcha, J. M., & Levy, W. B. (1996b). *A neural network solution to the transverse patterning problem depends on the repetition of the input code*. Manuscript submitted for publication.

---

Received August 22, 1996; accepted April 9, 1997.