Setting the Activity Level in Sparse Random Networks

Ali A. Minai
William B. Levy
Department of Neurosurgery, University of Virginia,
Charlottesville, VA 22908 USA

We investigate the dynamics of a class of recurrent random networks with sparse, asymmetric excitatory connectivity and global shunting inhibition mediated by a single interneuron. Using probabilistic arguments and a hyperbolic tangent approximation to the gaussian, we develop a simple method for setting the average level of firing activity in these networks. We demonstrate through simulations that our technique works well and extends to networks with more complicated inhibitory schemes. We are interested primarily in the CA3 region of the mammalian hippocampus, and the random networks investigated here are seen as modeling the a priori dynamics of activity in this region. In the presence of external stimuli, a suitable synaptic modification rule could shape this dynamics to perform temporal information processing tasks such as sequence completion and prediction.

1 Introduction

Recurrent networks of neural-like threshold elements with sparse, asymmetric connectivity are of considerable interest from the computational and biological perspectives. From the perspective of neurobiology, sparse, recurrent networks are especially useful for modeling the CA3 region of the mammalian hippocampus. Due to its recurrent connectivity, CA3 is thought to play a central role in associative memory (Marr 1971; Levy 1989; McNaughton and Nadel 1989; Rolls 1989; Treves and Rolls 1992) and the processing of temporal information (Levy 1988, 1989). In both cases, CA3 is seen as a system capable of learning associations between patterns and using these associations for spatiotemporal pattern recognition. Synaptic counts (Amaral et al. 1990) indicate that CA3 is a sparsely connected recurrent network. Another characteristic of the system is that the primary pyramidal cells are inhibited by a much smaller population of interneurons (Buzsáki and Eidelberg 1982). The relative scarcity of
inhibitory cells also implies that inhibition is broadly directed, with each interneuron inhibiting a large number of primary cells in its vicinity. Few neural network models take these two aspects of CA3 into account. Motivated to understand the role played by these characteristics in the network's dynamical behavior, we have recently investigated a class of sparse, random networks with nonspecifically directed inhibition (Minai and Levy 1993a,b) and have found fixed point, cyclical, and effectively aperiodic behavior. We have also developed a simple model relating the level of network activity to parameters such as inhibition, firing threshold, and the strength of excitatory synapses.

Several researchers have studied the dynamics of sparse, asymmetric networks within the framework of statistical mechanics (Derrida and Pomeau 1986; Sompolinsky and Kanter 1986; Derrida et al. 1987; Gutfréund and Mézard 1988; Kree and Zippelius 1991). One interesting conclusion to emerge from some studies is that, above some critical value, both sparseness (Kürtken, 1988) and asymmetry (Spitzner and Kinzel 1989a,b; Niitzel 1991) lead to an effectively aperiodic network dynamics, which may be called effectively aperiodic.

In this paper, we present a simplification of our model based on an approximation to the standard error function. This simplification allows accurate prediction of network activity using a closed form equation. The level of activity in a recurrent network is of crucial importance for learning. A low but consistent level of activity enables the network to recode its stimuli as sparse patterns, thus decreasing interference between the representations of different stimuli and increasing capacity. Our model demonstrates how a CA3-like network without synaptic modification can control its level of activity using inhibition.

2 Network Specification and Firing Probability

Our network model is qualitatively similar to the associative memory model proposed by Marr (1971) and later investigated by others (Gardner-Medwin 1976; McNaughton and Nadel 1989; Willshaw and Buckingham 1990; Gibson and Robinson 1992). A network consists of $n$ binary ($0/1$) primary neurons, each with identical firing threshold $\theta$. The network's connectivity is generated through a Bernoulli process. Each neuron $i$ has probability $p$ of receiving a fixed excitatory connection of strength $w$ from each neuron $j$ (including itself). The existence/nonexistence of such a connection is indicated by the $1/0$ random variable $c_{ij}$ where $P(c_{ij} = 1) = p$. Inhibition is provided by a single interneuron that takes input from all primary neurons and provides an identical shunting conductance proportional to its input to all primary neurons. Defining $K$ as the inhibitory weight and $m(t)$ as the number of active neurons at time $t$,
the excitation $y_i$ and output $z_i$ of neuron $i$ at time $t$ are given by

$$y_i(t) = \frac{w \sum_{j=1}^{n} c_{ij} z_i(t-1)}{w \sum_{j=1}^{n} c_{ij} z_i(t-1) + K \sum_{j=1}^{n} z_j(t-1)}$$

$$z_i(t) = \begin{cases} 1 & \text{if } y_i(t) \geq \theta \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

with $y_i(t) = 0$ for all $i$ if $m(t-1) = 0$, so that once the network becomes totally inactive, it cannot spontaneously reactivate itself.

Substituting equation 2.2 in equation 2.1 and defining $\alpha = K/(1-\theta)w$, the condition for firing is obtained as

$$\sum_{j=1}^{n} c_{ij} z_j(t-1) \geq \alpha m(t-1) \quad (2.3)$$

which means that, to fire at time $t$, neuron $i$ must have at least $\lceil\alpha m(t-1)\rceil$ active inputs, where $\lceil x \rceil$ denotes the smallest integer not less than $x$. Note that, due to the effect of averaging, the right-hand side of equation 2.3 is independent of $i$, and the inequality represents a universal firing condition. The firing condition also demonstrates that it is the composite, dimensionless parameter $\alpha$ that determines the dynamics of the network. In effect, $\alpha$ represents the relative strength of inhibition and excitation in the network, weighted appropriately by the threshold.

If $m(t-1) = M$, the average firing probability for a neuron $i$ at the next time step $t$ is

$$\text{average firing probability } \equiv \rho(M; n, p, \alpha)$$

$$= \sum_{k=\lceil\alpha M\rceil}^{M} \binom{M}{k} p^k (1-p)^{M-k} \quad (2.4)$$

If $M$ is sufficiently large, the binomial can be approximated by a gaussian, giving

$$\rho(M; n, p, \alpha) \approx \frac{1}{2} \left[ 1 - \text{erf} \left( \frac{X}{\sqrt{2}} \right) \right] \quad (2.5)$$

where

$$X \equiv \frac{\lceil\alpha M\rceil - Mp}{\sqrt{Mp(1-p)}}$$

A reasonable criterion for applying this approximation is $Mp > 5$ and $M(1-p) > 5$. In sparse networks with $p < 0.5$, only the first condition is relevant, and the criterion is $M > 5/p$. Assuming that neurons fire independently, as they will tend to do in large and sparsely connected
networks (Minai and Levy 1993a), we obtain a stochastic return map relating \( m(t) \) to \( m(t - 1) \):

\[
m(t) = n'p[m(t - 1)] + O(\sqrt{n})
\]

(2.6)

Thus, the expected activity at time \( t \) is

\[
\langle m(t) \mid m(t - 1) \rangle = n'p[m(t - 1)]
\]

(2.7)

In the long term, the activity is attracted to 0 or to an \( O(\sqrt{n}) \) region around \( \bar{m} \), the point satisfying the fixed-point condition: \( \bar{m} = n'p(\bar{m}) \).

Alternatively, one might look at the activity level \( r(t) \equiv n^{-1}m(t) \) rather than total activity \( m(t) \) to obtain the map

\[
r(t) = p[nr(t - 1)] + O(1/\sqrt{n})
\]

(2.8)

and expected activity level

\[
\langle r(t) \mid r(t - 1) \rangle = p[nr(t - 1)]
\]

(2.9)

Here \( r(t) \) is an instance of what Amari (1974) calls the macrostate of the system. The activity level fixed-point of the network is defined as \( \bar{r} \equiv n^{-1}\bar{m} \). The qualitative behavior of the network depends in large measure on the value of \( \bar{r} \), as discussed in our earlier studies (Minai and Levy 1993a,b).

3 Approximating the Firing Probability

While equations 2.4 and 2.5 can be used to obtain the activity map (and thus the activity level map) for a network with specified parameters, it is useful to look for a closed form, both for ease of calculation and for analytical manipulation. Such a closed form can be found using the approximation (see, e.g., Hertz et al. 1991)

\[
erf\left(\frac{x}{\sqrt{2}}\right) \approx \tanh\left(\sqrt{\frac{2}{\pi}}x\right)
\]

(3.1)

Substituting equation 3.1 in equation 2.5, we obtain

\[
\rho(M; n, p, \alpha) \approx \frac{1}{2} \left\{ 1 - \tanh \left[ \sqrt{\frac{2}{\pi}} \left( \frac{\alpha M - Mp}{\sqrt{Mp(1 - p)}} \right) \right] \right\}
\]

(3.2)

If \( M \) is large enough, \( [\alpha M] \approx \alpha M \), which simplifies equation 3.2 to

\[
\rho(M; n, p, \alpha) \approx \frac{1}{2} \left\{ 1 - \tanh \left( \frac{\sqrt{M}}{T} \right) \right\}
\]

(3.3)

where

\[
T = \frac{1}{\alpha - p} \sqrt{\frac{\pi p - (1 - p)}{2}}
\]

Equation 3.3 shows that the average firing probability is an increasing function of \( M \) for \( \alpha < p \), a decreasing one for \( \alpha > p \), and constant at 0.5 if \( \alpha = p \) (see Figure 1).
Figure 1: The average firing probability at time $t$ as a function of the activity at time $t-1$ for a 1000 neuron network with $p = 0.05$. Each curve corresponds to a different value of $\alpha$, as indicated on the graph. Equation 3.3 was used for the calculation. The point where the diagonal crosses each curve is the predicted stable activity level for the corresponding value of $\alpha$.

4 Setting Parameter Values to Obtain a Stable Activity Level

The most useful application of this model is in specifying the average level of activity in a network. Since the activity level stabilizes around $\bar{r}$ in the long term, we can use it as an estimate of the average activity level, $\langle r \rangle$, though, strictly speaking, there might be a slight discrepancy between the estimate and the actual value due to the asymmetric shape of the firing probability curve. As $r(t)$ becomes more confined with increasing network size, this discrepancy becomes less and less significant.
To obtain a specific $\bar{r}$ by setting $\alpha$, we just need to find the $\alpha$ satisfying the activity fixed point equation

$$\dot{m} = n\rho(m)$$

and substitute $\dot{m} = nr$. This gives

$$\alpha(\bar{r}) \approx p + \sqrt{\frac{\pi p(1 - p)}{2nr}} \tanh^{-1}(1 - 2r)$$

As long as $\bar{r}$ is not too close to 0 or 1, this is an adequate method for setting $\alpha$, as shown by the results in Section 6. Note that the useful range of $\alpha$ is bounded as $0 \leq \alpha \leq 1$; for values of $\alpha$ larger than 1, the firing condition (equation 2.3) shows that $\rho(M) = 0 \forall M$. When $\bar{r}$ is too low, $\dot{m}$ is not large enough to allow the gaussian approximation of equation 2.5, and the model breaks down. However, as $n$ grows, lower values of $\bar{r}$ come within the range of satisfactory approximation. Using the criterion $Mp > 5$ for the gaussian approximation, we conclude that equation 4.2 can be applied for $\bar{r} > 5/np$.

5 Extension to the Multiple Interneuron Case

So far in our model, we have assumed that inhibition is mediated by a single interneuron. In this simplification, we follow previous studies such as those by Gardner-Medwin (1976) and Gibson and Robinson (1992). However, as we argue in this section, the neuron model of equation 2.1 is also consistent with a population of interneurons, provided that these interneurons respond faster than the primary neurons and are statistically identical. Thus, let there be a set $I$ of $N$ interneurons, where each interneuron $I$ receives an input synapse of weight $u$ from each primary cell $j$ with a fixed probability $\gamma$ and projects back to each $j$ with probability $\lambda$ and weight $v$. Also, let $E$ denote the set of $n$ primary cells. The net excitation to each $I$ is given by

$$Y_I(t) = u \sum_{j \in E} c_{ij} z_j(t - 1)$$

where $c_{ij}$ is a binary variable indicating the presence or absence of a connection from $j$ to $I$. Since interneurons are postulated to be linear (see the Appendix), the output of $I$ is

$$Z_I(t) = CY_I(t)$$

where $C$ is a constant. Based on physiological evidence that hippocampal interneurons respond faster than pyramidal cells (Buzsáki and Eidelberg 1982; McNaughton and Nadel 1989), we assume that $I$ responds instantaneously to its input (unlike the primary cells that take one time unit to
The inhibitory input to a primary cell \( i \) at time \( t \) is then given by

\[
\phi_i(t) = v \sum_{i=1} c_{il} Z_i(t) = uv C \sum_{i=1} c_{il} \sum_{j \in E} c_{ij} Z_i(t-1)
\]

(5.3)

If \( \gamma m(t-1) \) is large enough, the distribution of \( Z_i(t) \) for each \( l \in I \) will approximate a gaussian with a mean value of \( u C \gamma m(t-1) \). By the central limit effect, then, \( \phi_i(t) \) will also be approximately normally distributed with mean \( uv C \gamma \lambda Nm(t-1) \). Thus, we can rewrite equation 2.1 as

\[
y_i(t) = \frac{w \sum_{j \in E} c_{ij} Z_j(t-1)}{w \sum_{j \in E} c_{ij} Z_j(t-1) + Km(t-1) + \eta(u, v, C, \gamma, \lambda, m(t-1), N)}
\]

(5.4)

where \( K \equiv uv C \gamma \lambda N \) and \( \eta(u, v, C, \gamma, \lambda, m(t-1), N) \) is a random fluctuation that is \( O(\sqrt{Nm(t-1)}) \) and becomes increasingly insignificant as \( Nm(t-1) \) increases. Thus, equation 2.1 represents a reasonable approximation for equation 5.4 in large networks with a few thousand primary cells and a few hundred inhibitory neurons—especially if \( \lambda \) is not too small. It should be noted that the multiple interneuron scheme described above is equivalent to having nonuniformly distributed random inhibitory connections between primary neurons, albeit with a shunting effect.

Calculations from intracellular studies of inhibition in CA3 (Miles 1990) suggest that each pyramidal cell receives inhibitory synapses of widely varying strengths from 10 to 50 interneurons in its general neighborhood. The same study indicates that a specific inhibitory interneuron makes synapses of roughly equal strength with a large proportion of pyramidal cells in its neighborhood. Together, these factors suggest that the inhibition to neighboring pyramidal cells is highly correlated in amplitude and phase (Miles 1990). Thus, our model does capture part of the inhibitory structure in CA3, though more realistic models will probably be needed as the physiology becomes clearer.

6 Results and Discussion

To show that the relationship between \( \alpha \) and \( \bar{r} \) as expressed in equation 4.2 can be used to estimate the average activity level, we simulated a number of 300 and 1000 neuron networks and obtained empirical data to compare against the model. We obtained the average activity level, \( \langle r \rangle \), by running each network for 2000 steps and averaging its activity level over the last 1000 of these steps. We simulated seven different, randomly generated networks for each value of \( \alpha \) and averaged the \( \langle r \rangle \)'s obtained in the seven cases. The results for \( n = 300 \) and \( n = 1000 \) are shown in Figure 2a and b, respectively. It is clear from the graphs that \( \bar{r} \), as calculated from equation 4.2, is a good estimator for \( \langle r \rangle \) when the activity level is not too high (\( > 0.8 \) or so) or too low (\( < 0.15 \) or so). One notable feature of the data is the large variance in the empirical average activity...
Figure 2: The predicted and empirical activity level fixed points for different values of $\alpha$ in (a) a 300 neuron network, and (b) a 1000 neuron network, with $p = 0.05$. The solid line shows the curve predicted by equation 4.2, while the bullets show the empirical values. The activity level is averaged over seven different networks for each $\alpha$ value. All runs were started from the same initial condition, and the last 1000 steps of a 2000 step simulation were used in each case. It is clear that the model works very well when $\bar{r}$ is not too low or too high. The large error bars at low activity levels are due to the fact that some networks switched off while others converged to low activity cycles.

level at high $\alpha$. This is due to the fact that some networks with these $\alpha$ values switched off in the first 1000 steps while others settled down to relatively short limit cycles of the activity level predicted by equation 4.2. This trivial network behavior was discussed in our earlier study (Minai and Levy 1993a). Suffice it to say that this phenomenon is mediated by very low-activity states, and we expect it to become less significant in
larger networks where even low-activity states have a large number of active neurons.

Figure 3 plots the empirically measured averaged activity level of a network with 1000 excitatory neurons and 50 inhibitory neurons for various values of $C$ with fixed $u, v, \gamma, \lambda$. The solid curve indicates the prediction generated using equation 4.2 with $K = uvC\gamma\lambda N$. The results at activity levels above 0.5 are not as good as in the single interneuron case, mainly because the variance in the inhibitory term is relatively high, but, comparing Figure 3 with Figure 2b, it is clear that the model of equation 4.2 works better in the multiple interneuron case when the activity level is low—presumably because the averaging in the inhibitory term makes simultaneous switch-off of all neurons less likely. Since the overall performance of the model should improve as $n$ and $N$ increase, there is reason to expect that equation 4.2 can be used to predict activity levels in large networks with multiple interneurons and very low activity.

7 Biological Considerations

The results given above show that equation 4.2 is a good model for relating average activity to $n$. However, the network model is supposed to be a representation of biological networks such as the CA3, and it is important to put it in its full biological context. Above all, it is necessary to demonstrate that the model can be applied to networks of realistic size without running into problems.

The most obvious potential problem, as implied by equation 4.2 and borne out by a comparison of Figure 2a and b, is that the dependence of $r$ on $\alpha$ tends toward a step function as $n$ increases. In infinitely large networks, therefore, all neurons fire if $\alpha$ is less than $p$ and none fire if it is greater than $p$, leading to an all-or-none activity situation. Even in large finite sized networks, however, obtaining a moderate activity level is problematic unless $\alpha$ is set very precisely. Since $\alpha$ is in arbitrary units, it is difficult to say exactly what degree of precision is prohibitive, but, as shown by equation 4.2, activity in large networks is easier to control at the lower end of the activity level spectrum. This is consistent with the physiological observation that activity levels in the rat CA3 are typically less than 1% (Thompson and Best 1989). Since Amaral et al. (1990) estimate that the rat CA3 has 300,000 or so pyramidal cells and a connectivity of around 1.9%, we calculate the relationship between $r$ and $\alpha$ for $n = 300,000$ and $p = 0.02$ as predicted by our model (Fig. 4). The activity levels shown are well above the value of 0.0008 needed for the gaussian approximation according to the criterion given earlier. It is apparent that $r$ varies smoothly with $\alpha$ in the range shown and $\alpha$ can, therefore, be used to control activity at typical CA3 levels.

In the biological context, one must also try to account for the effects of synaptic modification. Whereas our model treats all excitatory synapses
Figure 3: The predicted and empirical activity level fixed points for different values of $\alpha$ in a network with 1000 primary neurons and 50 inhibitory interneurons. The connectivity parameters are $p = 0.05$, $\gamma = 0.1$, and $\lambda = 0.5$. The firing threshold is set to 0.5 and the connection weights are $w = 1.0$, $u = 1.0$, and $v = 1.0$. The values of $w$ and $\theta$ mean that $\alpha = K$. The bullets indicate the empirically obtained average activities, while the solid line shows the curve predicted by equation 4.2 for a 1000 neuron network with a single interneuron and $\alpha$ calculated using $K = uvC\gamma\lambda N$. The activity level is averaged over seven different networks for each $\alpha$ value. All runs were started from the same initial condition, and the last 1000 steps of a 2000 step simulation were used in each case. The model works adequately for moderate activity values, and much better than in the single interneuron case for low activities. Its overall performance should improve when the number of neurons and interneurons is larger.
as identical, real synapses have varying strengths. Furthermore, due to the effects of synaptic modification, highly potentiated synapses on a primary cell would tend to be activated in correlated groups, not at random as in our model. This is both a cause and, through recurrence, an effect of correlated firing among the neurons themselves. The variation in synaptic strength and correlated input activity means that the excitation, \( w \sum_i c_{ij}z_j(t - 1) \), to a typical CA3 primary cell \( i \) will not necessarily have a unimodal gaussian distribution, as is the case in our model. This more
complex distribution of excitation to different cells will also alleviate the all-or-none activity problem described in the previous paragraph. Thus, it is best to see the random case treated in this paper as the starting point of the synaptic modification process, and the dynamics given by our model as the a priori or intrinsic dynamics of a CA3-like network. Synaptic modification can then be considered a "symmetry-breaking" process that gradually distorts the intrinsic dynamics into the spatiotemporal patterns implicit in the received environmental information. There have been some studies of synaptic modification in networks similar to ours, but only in the context of associative memory (Marr 1971; Gardner-Medwin 1976; Palm 1980; Gibson and Robinson 1992). Our main interest lies in the temporal aspects of network behavior and not just in the stability properties required by associative memory.

Finally, we turn to another interesting aspect of our model: the assumption of continuous-valued (linear) inhibition. There is some evidence that the response of inhibitory interneurons in the hippocampus differs from that of the primary cells in more than just its speed. Interneurons have very low response thresholds and respond with multiple spikes whose number (and onset latency) is directly related to the intensity of the stimulus to the interneuron (Buzsáki and Eidelberg 1982). When integrated postsynaptically, these spike trains could represent an almost continuous-valued signal with a significant dynamic range. Of course, this does not imply that the interneuron's response is linear in its stimulus as we, and others, have assumed. However, the analysis developed in this paper can be extended to some kinds of nonlinear inhibitory schemes, as described in the Appendix.

8 Conclusion

The mammalian hippocampus is a complex system and is probably involved in highly abstract information processing tasks such as recoding, data fusion, and prediction (Levy 1985, 1989; Rolls 1989; Rolls and Treves 1990; McNaughton and Nadel 1989). With its recurrent connectivity, it is natural to expect that the CA3 region plays an important role in any temporal processing done by the hippocampus (Levy 1989). In this paper, we have studied a recurrent network with CA3-like characteristics and have presented a model for relating its average activity level to parameters such as neuron firing threshold and the strength of inhibition. Using simulations, we have demonstrated that this model successfully relates network parameters to the activity level over a reasonable range, and that it is easily extended to situations with multiple inhibitory interneurons. Our model thus provides an understanding of the intrinsic dynamics of the untrained random network, and from this we can proceed to the problem of temporal learning and prediction through synaptic modification.
Appendix: Nonlinear Inhibition

We could rewrite equation 5.2 more generally as $Z_i(t) = g[m(t-1)]$, where $g(x)$ is an appropriate monotonically nondecreasing function (e.g., a sigmoid, as is often the case in artificial neural networks). Furthermore, since there is evidence of spontaneous, low-intensity firing in interneurons (Buzsáki and Eidelberg 1982), one could postulate an inhibition of the general form $Kg[m(t-1)] + \kappa$, where $\kappa$ is a small constant offset. The analysis given in the paper for linear inhibition transfers directly to the nonlinear case with offset. Thus, equation 2.1 becomes

$$y_i(t) = \frac{w \sum_{j=1}^{n} c_{ij}z_j(t-1)}{w \sum_{j=1}^{n} c_{ij}z_j(t-1) + Kg[m(t-1)] + \kappa}$$  \hspace{1cm} (A.1)

Following the logic used in the linear case, and applying the hyperbolic tangent approximation to the error function, we get

$$\rho(M; n, p, \alpha) \approx \frac{1}{2} \left[ 1 - \tanh \left( \frac{2 \alpha g(M) + \beta - Mp}{\sqrt{\pi}} \sqrt{Mp(1-p)} \right) \right]$$  \hspace{1cm} (A.2)

where $m(t-1) = M$ and $\beta \equiv \theta \kappa/(1-\theta)w$. As before, we have taken $[\alpha g(M)] \approx \alpha g(M)$. This, in turn, leads to a relationship between $\alpha$ and $\alpha$, albeit a slightly more complicated one than in the linear case:

$$\alpha(\tilde{r}) \approx \frac{nrp - \beta}{g(nr)} + \frac{\pi nrp(1-p)}{2g^2(nr)} \tanh^{-1}(1 - 2\tilde{r})$$  \hspace{1cm} (A.3)

The extension of this more general case to multiple, statistically identical interneurons is straightforward, and leads essentially to equations A.2 and A.3 for large networks if $g(x)$ is reasonably well-behaved.

Acknowledgments

This research was supported by NIMH MH00622 and NIMH MH48161 to W.B.L., and by the Department of Neurosurgery, University of Virginia, Dr. John A. Jane, Chairman. The authors would like to thank Dawn Adelsberger-Mangan for her constructive comments. The paper has also benefited greatly from the suggestions of the reviewers.

References


Received November 12, 1992; accepted May 26, 1993.