# Self-sustained activity in networks of gain-modulated neurons

Emilio Salinas [1]

*Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA*

**Abstract**

Simulation studies have shown that recurrently connected neurons are capable of sustaining non-uniform profiles of activity in the absence of tuned input. These attractor networks are the basis for models of working memory and other processes where information about transient stimuli is stored temporarily. In addition, there is strong evidence that neurons often interact by affecting each other's gain. Here I study a minimal recurrent network that takes gain interactions into account. I show analytically that, in agreement with results of computer simulations, a center-surround organization gives rise to two types of stable solutions: a uniform state in which all neurons fire at the same rate, and a self-sustained profile of activity that may be centered at any point in the network. This theoretical framework based on nonlinear neuronal interactions is, in general, a powerful tool for investigating recurrent network dynamics.

*Key words:* Gain fields, Divisive inhibition, Working memory, Neuron models.

## 1 Introduction

Models of mnemonic activity are typically based on recurrent networks with so-called line attractors or bump attractors [10,7,9]. These are networks in which many states or firing rate distributions are possible. These states are stable, so they can be maintained for relatively long times. Which state is observed depends on initial conditions or on additional inputs that steer the network from one state to another. Such models have been successful at describing working memory processes in prefrontal cortex [2], responses

[1] Supported by startup funds from WFUSM.
*E-mail address:* esalinas@wfubmc.edu.

of the head direction system [10] and responses of hippocampal place cells [6], among others [9]. However, analytical treatment of such attractors — particularly those that generate unimodal profiles of activity — has been rather limited [1,10,4]. Here I provide an analysis in which the attractor solution can be calculated explicitly in a straightforward way. The results show: (1) that the attractor solution may coexist with a state of uniform, low firing, (2) that the background input level has a strong influence on the the steady-state solutions, and (3) that the mechanisms for generating persistent responses and traveling waves of activity are very similar.

## 2 Nonlinear Network Equations

Consider a network of $N$ neurons where the response of unit $i$ is given by

$$\tau \frac{dr_i}{dt} = -r_i + \frac{A\,H\left(\sum_j w_{ij}\,r_j - \theta\right) \,+\, h}{s + v\sum_j r_j^2}\,, \tag{1}$$

where $H$ is the Heaviside step function ($H(x)=1$ if $x>0$ and is 0 otherwise), $\tau$ is a time constant, $A$ and $s$ are constants, and $h$ represents an excitatory external input or background activity that comes from outside the network.

Neurons in the above equations have a hard activation threshold $\theta$, as in Hopfield-type networks [3]. Crucially, however, they also incorporate gain interactions, for which there is abundant experimental evidence [5]. This is implemented through a mechanism similar to divisive normalization, which has served as an accurate empirical description of various experiments [8]. Thus, a neighboring neuron $j$ can either directly drive neuron $i$ to fire, through synaptic connection $w_{ij}$, or decrease its gain, through a synaptic connection of strength $v$. For simplicity, all gain interactions have equal strength, and no distinction is made between excitatory and inhibitory neurons.

The dynamics of this system depends on the network architecture, which is determined by the connectivity matrix $\mathbf{w}$. Here I'll study a connectivity profile equal to a square window; that is, $w_{ij}=w_{max}$ whenever $|i-j|\le N_w$ and $w_{ij}=0$ otherwise. Thus, each neuron excites equally $2N_w$ of its nearest neighbors. Periodic boundary conditions are used, so the array of neurons may be thought of as a ring [1,10,2]. Although all entries $w_{ij}$ are positive or zero, the organization is center-surround because the divisive mechanism effectively makes long-range interactions inhibitory.

Next, the firing rates of this system at steady state will be determined.

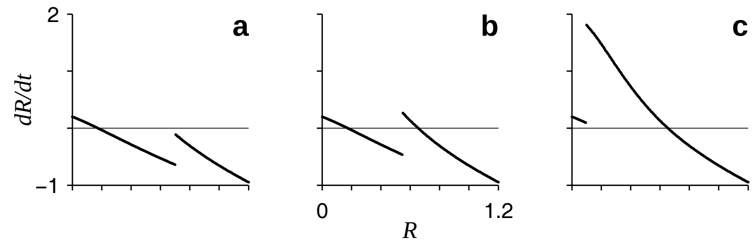


Fig. 1. Uniform rate solution. Plots show time derivative of $R$ vs $R$. Points at which the curves intersect the horizontal line are stable steady-state values of $R$. **a**, With $\theta = 2.10$ there is a single, low-rate steady state. **b**, With $\theta = 1.65$ there are two stable solutions. **c**, With $\theta = 0.3$ there is a single, high-rate steady state. Curves were obtained from Equation 2 with: $h = 0.11$, $A = 1$, $w_{tot} = 3$, $N = 100$, $v = 0.027$, $s = 0.56$, $\tau = 1$.

## 3 Uniform-Rate Solution

The first solution that can be obtained is one in which all units fire at the same rate, denoted as $R$. To see this, set $r_i = R$ in Equation 1,

$$\tau \frac{dR}{dt} = -R + \frac{A\, H(w_{tot}\, R - \theta) \;+\; h}{s + vNR^2}\,. \tag{2}$$

Here $w_{tot}$ is the total synaptic strength upon a postsynaptic neuron; that is, $w_{tot} = \sum_j w_{ij} = 2\, w_{max}\, N_w$. At steady state all derivatives are equal to zero. For the above expression there are two possible steady-state values, a low and a high one. This can be seen graphically in Figure 1. The low firing level occurs when $R\, w_{tot} < \theta$ and must satisfy $R = h/(s + vNR^2)$; the high firing level occurs when $R\, w_{tot} > \theta$ and must satisfy $R = (A + h)/(s + vNR^2)$.

Note that the uniform-rate solution applies regardless of the actual connectivity pattern, because it only depends on the total synaptic strength being the same for all units. This is a normalization condition on the matrix $\mathbf{w}$. As long as this condition is satisfied and the rest of the parameters are also equal across neurons, this solution will exist.

## 4 Self-Sustained Activity

The second solution to Equation 1 occurs when the firing rates in the network form a self-sustained profile of activity. The shape of this profile is the same as for the connections: a square window (Figure 2a, right). The reason is this. Suppose that initially the profile of activity is a square win-

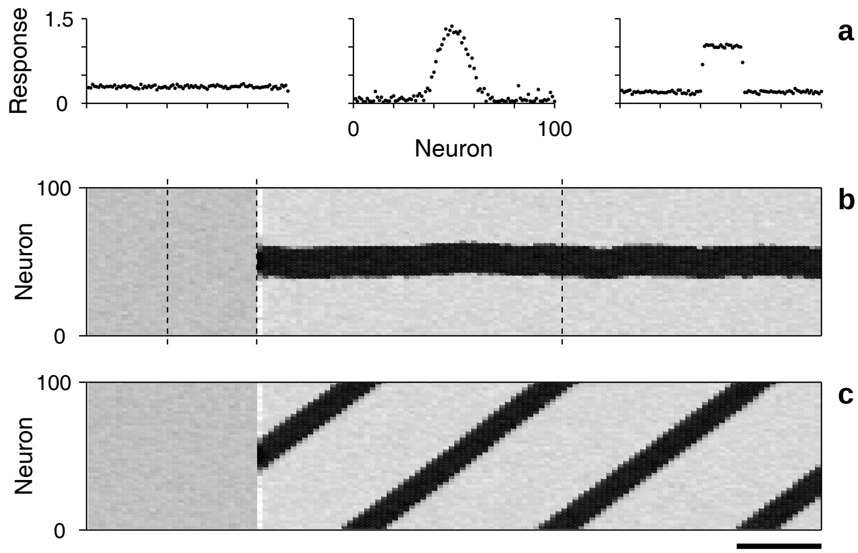


Fig. 2. Self-sustained activity. **a**, Responses of all neurons at the three points in time indicated by dashed lines in the panel below. **b**, Activity of all neurons as a function of time. Darkness level indicates response intensity. Initially all neurons fire at a low, uniform rate. A transient input (middle dashed line) switches the network to a square profile of activity. **c**, As in *b*, but with a connectivity pattern shifted by 1 unit. Scalebar equals 30 time units. Here $N_w = 15$, $w_{max} = 0.1$, $h = 0.25$, $s = 0.63$, $\theta = 1.8$, $\Delta t = 0.1$; other parameters as in Figure 1.

dow centered at some point. The operation $\sum_j w_{ij}\, r_j$ then corresponds to a convolution of two square functions, which results in a function that looks like a trapezoid. The step function $H$ takes this trapezoid minus the threshold $\theta$ as its argument, which results in another square function centered at the same point as the original. The divisive normalization mechanism determines the width, height and baseline of the activity profile, but not its shape or location. The width, height and baseline can be calculated analytically as follows.

Consider a stationary profile of activity that has a width $2N_r$. In this state, $2N_r$ neurons fire at a high rate $r_{max}$ and $N - 2N_r$ neurons fire at a baseline rate $r_{min}$. From Equation 1, at steady state these rates must satisfy

$$r_{min} = \frac{h}{s + v2N_r\, r_{max}^2 + v(N - 2N_r)r_{min}^2} \tag{3}$$

$$r_{max} = \frac{A + h}{s + v2N_r\, r_{max}^2 + v(N - 2N_r)r_{min}^2}\,. \tag{4}$$

Three main results follow from this:

$$
\begin{aligned}
r_{min} &= r_{max}\,\frac{h}{A+h} && (5)\\
N_r &= \frac{A+h-s\,r_{max}-v\,N\,r_{max}\,r_{min}^2}{2\,v\,r_{max}\,(r_{max}^2-r_{min}^2)} && (6)\\
r_{max} &= \frac{\theta}{N_w w_{max}}\,\frac{A+h}{A+2h}\,. && (7)
\end{aligned}
$$

The first result is obtained by dividing Equations 3 and 4 by each other. The second result is the same as Equation 4 but having solved for $N_r$. The third expression results from this observation: the difference between Equations 3 and 4 is whether the argument of the step function in Equation 1 is larger or smaller than zero, so for a neuron at the edge of the square window the argument must be equal to zero. Hence, for such edge unit $\sum_j w_{ij}\,r_j = N_w(r_{min}+r_{max}) = \theta$; the second equality and Equation 5 lead to Equation 7. The first equality requires that $N_w$ be smaller than $2N_r$, but by analyzing the output of $H$ applied to the trapezoid function it can be seen that $N_r > N_w/2$, so the condition is always satisfied. In this way, $r_{max}$, $r_{min}$ and $N_r$ are obtained in terms of the network parameters.

Figures 2a, 2b show an example of the behavior of the network generated by simulating Equation 1 with $N = 100$ neurons. To test for robustness, Gaussian noise ($\sigma = 0.02$) was added to all responses in each time step. Initially all neurons fire at a low level $R = 0.3$; then the responses around neuron 50 are increased transiently, after which a square pattern of activity develops. Its parameters are: $r_{min} = 0.2$, $r_{max} = 1$ and $N_r = 10$, which match the analytic calculations (for network parameters, see caption). Note that Figure 2 shows only one example, but the self-sustained activity profile can indeed be centered anywhere within the array.

## 5 Discussion

In spite of its simplicity, the network model presented here displays interesting dynamic behaviors that can be studied analytically and that, qualitatively, match the activity of much more complex models with persistent activity [2,4]. The model also provides valuable insight into how the various network parameters influence output activity. For instance, the strength of inhibitory interactions $v$ affects only the width of the square profile. The width also depends strongly on the background input $h$, although $h$ influences $r_{min}$, $r_{max}$, and the uniform rates as well. Because $h$ corresponds to input from external sources, it may change dynamically, and may hence be used for turning the network on or off. For instance, lowering $h$ to 0.16 in Figure 2 completely abolishes the self-sustained activity. Therefore, the uniform background may act as a control signal on the network.

Finally, an interesting extension of the results is the case of a square pattern of connections that is slightly shifted. That is, $w_{ij} = w_{max}$ if $|i - j - \Delta| \leq N_w$ and $w_{ij} = 0$ otherwise. By applying $H$ to the convolution between an initial square profile of activity and the shifted connections, it can be seen that the shift $\Delta$ does not change the shape of the profile; however, it moves it to a location $\Delta$ units away. Thus, the asymmetry generates a traveling wave. Figure 2b shows an example. It is not difficult to show that the speed of the wave should be proportional to $\Delta$, which is indeed what is found in simulations. At any instant, the activity profile is approximately square, with the same parameters as the stationary profile. This suggests that the problem of generating a stable, self-sustained pattern of activity is very similar to the problem of generating a stable, propagating wave of activity.

In conclusion, the present framework may be a powerful tool for analyzing the dynamics of recurrent neural networks.

## References


[1] R. Ben-Yishai, R.L. Bar-Or RL and H. Sompolinsky, Theory of orientation tuning in visual cortex. Proc. Natl. Acad. Sci. USA 92 (1995) 3844–3848.

[2] A. Compte, N. Brunel, P.S. Goldman-Rakic and X.-J. Wang, Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb. Cortex 10 (2000) 910–23.

[3] J. Hertz, A. Krogh and R.G. Palmer, Introduction to the theory of neural computation. (Addison-Wesley, New York, 1991).

[4] C.R. Laing and C.C. Chow, Stationary bumps in networks of spiking neurons. Neural Comput. 13 (2001) 1473–1494.

[5] E. Salinas E and P. Thier, Gain modulation: a major computational principle of the central nervous system. Neuron 27 (2000) 15–21.

[6] A. Samsonovich A and B.L. McNaughton, Path integration and cognitive mapping in a continuous attractor neural network model. J. Neurosci. 17 (1997) 5900—5920.

[7] H.S. Seung, How the brain keeps the eyes still. Proc. Natl. Acad. Sci. USA 93 (1996) 13339–13344.

[8] O. Schwartz and E.P. Simoncelli, Natural signal statistics and sensory gain control. Nat. Neurosci 4 (2001) 819–825.

[9] X.-J. Wang, Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci. 24 (2001) 455–463.

[10] K. Zhang, Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. J. Neurosci. 16 (1996) 2112–2126.