Self-sustained activity in networks of gain-modulated neurons

Emilio Salinas

Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, NC 27157
esalinas@wfubmc.edu

Abstract

Simulation studies of various network models have shown that recurrently connected neurons are capable of exhibiting active states that are sustained in the absence of any external input. These attractor networks are the basis for models of working memory and other processes where information about transient stimuli is stored termporarily. Using a novel theoretical framework to describe network interactions, I prove analytically that a network with sufficiently strong recurrent connections generates a self-sustained Gaussian profile of activity. I study the conditions under which this solution exists, and show that they are extremely general, in agreement with the results of computer simulations.

Introduction

Models that include some kind of mnemonic activity are typically based on so-called line attractors or bump attractors (Cannon et al., 1983; Ben-Yishai et al., 1995; Zhang, 1996; Salinas and Abbott, 1996; Seung, 1996; Hahnloser et al., 1999). These are networks in which many states or firing rate distributions are possible. These states are stable, so they can be maintained for a long time. 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 quite successful at describing working memory processes in prefrontal cortex (Compte et al., 2000), responses of the head direction system (Zhang, 1996) and responses of hippocampal place cells (Samsonovich and McNaughton, 1997), among others. However, analytical treatment of such attractors has been rather limited (Ben-Yishai et al., 1995; Zhang, 1996). Here I provide an analysis in which the attractor solution can be calculated explicitly in a straightforward way.

Nonlinear Network Equations and Their Solutions

The network equations I use incorporate gain control using a mechanism similar to divisive normalization (Simoncelli and Heeger, 1998; Schwartz and Simoncelli, 2001). Consider the following system of N interacting neurons,

$$\frac{dR_i}{dt} = -R_i + B + \frac{\left(h_i + \sum_j W_{ij} R_j\right)^2}{s + \left(v \sum_j R_j\right)^2}, \qquad (1)$$

where R_i is the firing rate of neuron i. Here the neurons are driven through synaptic weights $\boldsymbol{W}$, and the sum in the denominator represents the gain regulation mechanism. For simplicity, all connection weights are positive and all gain connections are equal to v. The term h_i represents an external input to that comes from outside the network and is not affected by its activity. B is also constant and serves to set the background rate in the simulations.

To proceed, first consider a relatively large number of neurons, so that sums over cells can be replaced by integrals. With this formulation, the vector of firing rates turns into a continuous function $R(a)$, where a acts as the neuron's index; similarly, the connection matrix $\boldsymbol{W}$ turns into a function $W(a,b)$ of two variables. To model a center-surround connectivity, the direct connections have a Gaussian shape with standard deviation σ,

$$W(a,b) = W(a-b) = w \exp\left(-\frac{(a-b)^2}{2\sigma^2}\right). \tag{2}$$

Thus, setting B and all external inputs to zero, the steady-state solution of Equation 1 in this case becomes

$$R(a) - \frac{\left(\rho \int W(a-b)\, R(b)\, db\right)^2}{s + \left(v\, \rho \int R(b)\, db\right)^2} = 0\,, \tag{3}$$

where ρ is the density of neurons; this factor ensures that the sum over neurons Σ_j equals the integral $\int db$. This equation has two types of solution. The first is a constant, so that all neurons have the same firing rate R_c. This rate satisfies

$$R_c - \frac{\rho^2\, w^2\, 2\pi\, \sigma^2\, R_c^2}{s + \rho^2\, v^2\, R_c^2} = 0\,, \tag{4}$$

This expression admits R_c=0 and two other possible values. The zero rate solution is always stable; the others may be real or imaginary, but this and their stability depend on the parameters.

The second solution to Equation 3 occurs when the firing rates in the network form a Gaussian profile of activity. That is, when

$$R(a) = R(a-x) = R_{\max}\ \exp\left(-\frac{(a-x)^2}{2\sigma^2}\right). \tag{5}$$

Here R depends on the difference a-x, where x is arbitrary and indicates the point of maximum activity. This means that the peak of the activity profile can be located anywhere within the network. To see that the above expression is indeed a solution of Equation 3, substitute it into the equation and calculate the two integrals, assuming that the limits are large enough so that edge effects can be ignored. There are two key elements to this result. First, the integral in the numerator is a convolution of two Gaussians, and the result of this operation is another Gaussian that is wider than any of the original two. However, the squaring operation then decreases the width, and the end result is a Gaussian with the same standard deviation σ as the one that entered into the integral. The second crucial component is the integral in the denominator, which is proportional to R_{max}. This gives rise to three possible values for R_{max}: zero, which is stable, and two more, which may be unstable and stable, respectively (there are other alternatives, but this is the relevant case). The stable, nonzero fixed point for R_{max} produces a stable profile of activity with a Gaussian shape. The unstable fixed point serves as a transition: above it, the Gaussian profile grows in amplitude until it reaches the stable solution; below it the profile shrinks until it reaches the stable constant solution R_c.

These results are confirmed through simulations, which are illustrated in Figure 1. Each neuron in the simulated network excites its neighbors, as specified by Equation 2, and follows the dynamics of Equation 1. The simulated network behaves exactly as predicted by the analysis. When the rates are initialized to values below a certain transition point, activity relaxes to a low rate that is identical for all units. In contrast, when the rates are initialized with a broadly peaked distribution above the transition point, a stable profile of activity with a single peak develops and stays stable. This peak can be located anywhere along the array, depending on the initial conditions.

Generalizations and implications

Interestingly, this result does not depend very much on the exact form of Equation 1. For instance, if the exponents in the numerator and denominator of Equation 1 are set to m and n, it can be shown that a stable, self-sustained Gaussian profile will still be obtained as long as $n>1$ and $m>n-1$.

Self-sustained activity can still be obtained even if the functional form of the interaction changes in other ways. For instance, if the connectivity pattern is square rather than Gaussian, and instead of a squaring nonlinearity in the numerator of Equation 1 a step function is applied, the system will still exhibit a behavior very similar to that shown in Figure 1. The difference is that resulting profile of activity will also have a square shape. This can also be proven rigorously.

In conclusion, diverse kinds of connectivity patterns and nonlinearities give rise to stable, self-sustained activity profiles. The recipe to obtain these dynamics seems to involve a combination of recurrent activity with some form of nonlinearity in the response, plus some form of gain control.

References

Ben-Yishai R, Bar-Or RL, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. PNAS 92:3844--3848.

Cannon SC, Robinson DA, Shamma S (1983) A proposed neural network for the integrator of the oculomotor system. Biol Cybern 49:127--136.

Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb Cortex 10:910--23.

Hahnloser R, Douglas RJ, Mahowald M, Hepp K (1999) Feedback interactions between neuronal pointers and maps for attentional processing. Nature Neurosci 2:746--752.

Salinas E, Abbott LF (1996) A model of multiplicative neural responses in parietal cortex. Proc Natl Acad Sci USA 93:11956--11961.

Samsonovich A, McNaughton BL (1997) Path integration and cognitive mapping in a continuous attractor neural network model. J Neurosci 17:5900--5920.

Schwartz O, Simoncelli EP (2001) Natural signal statistics and sensory gain control. Nat Neurosci 4:819—825.

Seung HS (1996) How the brain keeps the eyes still. Proc Natl Acad Sci USA 93:13339--13344.

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. Vision Res 38:743--761.

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

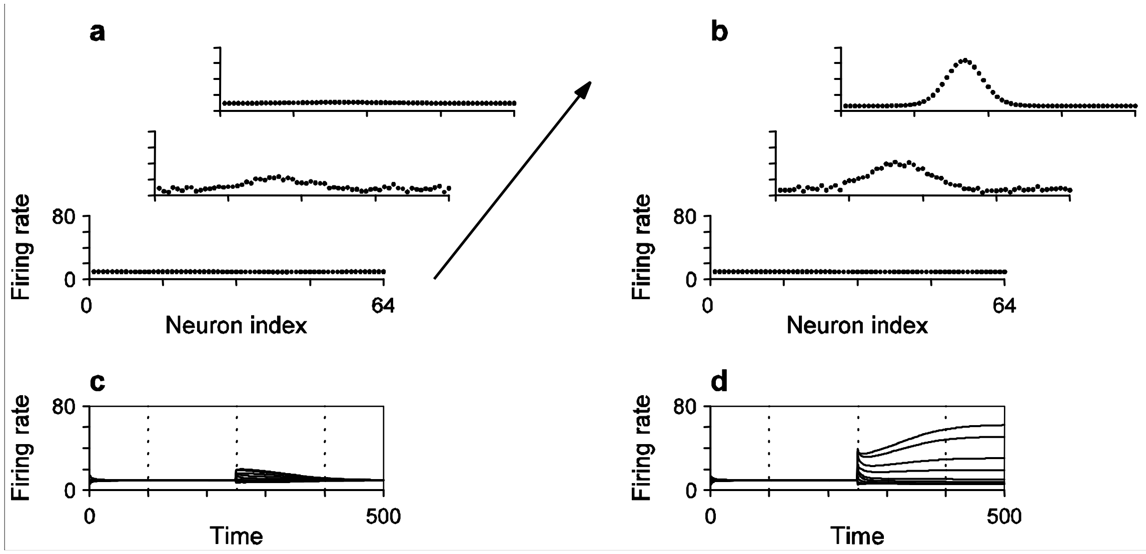

Figure 1. Responses in a gain-modulated network with Gaussian center-surround connectivity. Sixty four neurons were coupled using Equation 2 with w=0.0417, σ=3.2, and v=0.0021. **a, b,** Activities of all neurons at three points in time. **c, d,** Activity as a function of time for 9 of the neurons. Dotted lines mark the three points in time at which all responses are shown; long arrow indicates order. At t=0 all rates were set to random values around the low steady state. At t=250 the rates were shifted to higher values. In **a** and **c** the network relaxes back to the low steady state, around 10 spikes/s. In **b** and **d** the shift is above the transition point and the network switches to the high steady state, forming a Gaussian profile of activity. The profile is maintained without further input. Other parameters were: B=5, s=0.2846, ρ=1, N=64. Self-connections were eliminated and periodic boundary conditions were used.