ELSEVIER

Neurocomputing 38–40 (2001) 581–586

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Beyond bumps: Spiking networks that store sets of functions[☆]

Chris Eliasmith[a,*], Charles H. Anderson[b]

[a]Department of Philosophy, University of Waterloo, Waterloo, Ont, N2L 3G1 Canada
[b]Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO 63110, USA

Abstract

There are currently a number of models that use spiking neurons in recurrent networks to encode a stable Gaussian 'bump' of activation. These models successfully capture some behaviors of various neural systems (e.g., storing a single spatial location in working memory). However, they are limited to encoding single bumps of uniform height. We extend this previous work by showing how to construct and analyze realistic spiking networks that encode multiple 'bumps' of different heights. Our networks capture additional experimentally observed behavior (e.g., storing multiple spatial locations at the same time and the sensitivity of working memory to non-spatial parameters). © 2001 Published by Elsevier Science B.V.

Keywords: Function attractors; Neural representation; Working memory; Bump networks

1. Introduction

Stable Gaussian-shaped neuronal activities across a population of recurrently connected neurons without external perturbation, or stable 'bumps', have been successfully modeled by a number of researchers [7,3,2]. Bumps have been thought to be present in various neural systems including the head direction system [12], frontal working memory systems [7], parietal reach memory systems [11], and feature

☆This research has been supported by the McDonnell Center for Higher Brain Function at Washington University and the Mathers Foundation.

* Corresponding author.
E-mail addresses: celiasmith@uwaterloo.ca (C. Eliasmith), cha@shifter.wustl.edu (C.H. Anderson).

0925-2312/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S0925-2312(01)00419-2

selective visual systems [6]. Many of these systems can store functions more complicated than a simple uniformly-sized Gaussian bump. For example, there is evidence that parietal areas can hold multiple saccade targets in memory at the same time, suggesting that a multi-modal function is stored [8]. As well, it has been shown that the activity in parietal areas is sensitive to non-spatial parameters (e.g., shape [9] and intensity [10]), suggesting that bumps of activity at a single location should be of different heights under different stimulus conditions. None of the models proposed to date support either multiple bumps or bumps of different amplitudes.

In this paper, we extend previous work on single bump networks by showing how to construct and analyze realistic spiking networks that can encode multiple bumps of arbitrary amplitude. We begin our analysis with a one-dimensional network of rate-modeled neurons. We then show how it is possible to generate useful analytical results about this simple network. We discuss important extensions to the model, including how to implement the model in a spiking network (using leaky integrate-and-fire neurons), and how to construct higher dimensional models. Notably, the approach we employ is a general one which can be applied to constructing and analyzing many different kinds of neural circuits [4,5].

2. The simple rate model

The procedure outlined here is appropriate for constructing a function attractor for any set of functions. In this case, however, we restrict the set of functions to include all those functions made up of independent Gaussian-shaped functions of arbitrary height. We begin by defining the space in which the functions of interest, $f(x;\mathbf{A})$, reside

$$f(x;\mathbf{A}) = \sum_{n=1}^{D} A_n \chi_n(x). \tag{1}$$

$$A_n = \langle \chi_n(x) f(x;\mathbf{A}) \rangle_x, \tag{2}$$

where $\langle \chi_n(x) f(x;\mathbf{A}) \rangle_x = \int \chi_n(x) f(x;\mathbf{A})\,\mathrm{d}x$.

We choose a finite number, D, of orthonormal basis functions, $\chi_n(x)$, to define this functional space. In our simulations, the functional space is constrained to have low spatial frequencies, resulting in smooth functions. This is appropriate since we are interested in representing Gaussian bumps of arbitrary heights at any position.

Next, we assume that the neurons form a highly overcomplete representation of this same functional space. The encoding functions for this representation are taken to be the neuron response functions

$$a_i(\mathbf{A}) = G_i[\alpha_i \langle \tilde{\phi}_i(x) f(x;\mathbf{A}) \rangle_x + \beta_i]. \tag{3}$$

Assuming linear decoding, the estimate of the original function given this encoding is

$$\hat{f}(x;\mathbf{A}) = \sum_{i=1}^{N} [a_i(\mathbf{A}) + \eta_i] \phi_i(x). \tag{4}$$

The $a_i(\mathbf{A})$ are the neuron firing rates that represent the function $f(x;\mathbf{A})$ under some independent Gaussian distributed noise with a mean of zero, η_i. The firing rates themselves are determined by the neuron's encoding function, $\tilde{\phi}_i(x)$, a gain, α_i, a bias, β_i, and the neuron's rate function, G_i (e.g. a leaky integrate-and-fire response function). The encoding function in our case is a Gaussian, since the tuning curves found in the area we are simulating (parietal cortex) are Gaussian. The encoding function is used as a measure of how similar the represented function is to the preferred function for that neuron.

We can now find the decoding functions in the neuron space, $\phi_i(x)$, that provide our estimate, $\hat{f}(x;\mathbf{A})$, by minimizing the error between the original function (represented in the orthonormal space of $\chi_n(x)$) and that estimate (in the neuron space)

$$E = \sum_n MSE_n = \left\langle \int \left[\sum_n A_n \chi_n(x) - \sum_i [a_i(\mathbf{A}) + \eta_i] \phi_i(x) \right]^2 \mathrm{d}x \right\rangle_{\mathbf{A},\eta} \tag{5}$$

$$= \sum_n \left\langle \left[A_n^2 - \sum_i k_{in}^2 [a_i(\mathbf{A}) + \eta_i]^2 \right] \right\rangle_{\mathbf{A},\eta}. \tag{6}$$

In Eq. (6) we have expressed the neuron decoding functions as a linear sum of orthonormal basis functions, $\phi_i(x) = \sum_n k_{in} \chi_n(x)$, in order to ensure that they encode, at best, the same functional space as the orthonormal basis.

Minimizing this error with respect to k_{in} gives:

$$\mathbf{k} = \Gamma^{-1} \Upsilon, \tag{7}$$

where

$$\Gamma_{ij} = \langle a_i(\mathbf{A}) a_j(\mathbf{A}) \rangle_{\mathbf{A}} + \sigma_\eta^2 \delta_{ij},$$

$$\Upsilon_{in} = \langle a_i(\mathbf{A}) A_n \rangle_{\mathbf{A}}.$$

In order to *store* the set of functions $f(x;\mathbf{A})$, we must ensure that the dynamics of our network are stable. To determine how stable we expect the dynamics to be, we can explicitly write our expression for our optimal decoding coefficients from Eq. (7) as

$$k_{in} = \sum_j \Gamma_{ij}^{-1} \langle a_i(\mathbf{A}) A_n \rangle_{\mathbf{A}}. \tag{8}$$

Substituting this into our expression for mean square error (6), gives

$$MSE_n = \langle A_n^2 \rangle_{\mathbf{A}} - \sum_{ij} \langle a_i(\mathbf{A}) A_n \rangle_{\mathbf{A}} \Gamma_{ij}^{-1} \langle a_j(\mathbf{A}) A_n \rangle_{\mathbf{A}}.$$

Our representation of each element, A_n, of the $\mathbf{A}$ vector (which determines the signal) is accurate to within $\sqrt{MSE_n}$. Since we can reduce the *MSE* by increasing the number of neurons in the population, we can repeatedly encode and decode the signal with an error that can be made arbitrarily small given sufficient number of neurons.

Now we want to find weights in the neuron space that force the network to display this behavior. In particular, we want the network to decode the same function it encodes from time step to time step. The weights needed to give stable dynamics can be found as follows:

$$a_i^{m+1} = G_i[\alpha_i \langle \tilde{\phi}_i(x) f(x;\mathbf{A}^m)\rangle_x + \beta_i]$$

$$= G_i\left[\alpha_i \left\langle \tilde{\phi}_i(x) \sum_j [a_j^m + \eta_j]\phi_j(x)\right\rangle_{x,\eta} + \beta_i\right] \quad (9)$$

$$= G_i\left[\alpha_i \sum_j \omega_{ij} a_j^m + \beta_i\right]. \quad (10)$$

where

$$\omega_{ij} = \langle \tilde{\phi}_i(x)\phi_j(x)\rangle_x = \sum_n k_{jn}\langle \tilde{\phi}_i(x)\chi_n(x)\rangle_x.$$

The weights, in this case, are the projection of the encoding functions onto the decoding functions. These weights now define a set of attractors in the function space we defined earlier. These weights implement this kind of repeated encoding and decoding. The result is a set of function attractors with stable representations of the functions defined by, $f(x;\mathbf{A})$. Our simulations have verified this result, which shows that the amplitudes, A_n, are dynamically re-mapped to themselves in spite of the non-linear neuronal encoding procedure. The preceding procedure results in a localized connectivity matrix that supports stable dynamics for the set of functions determined by $\mathbf{A}$ as long as the gain after encoding and decoding is smoothly reduced from unity for spatial frequencies above the cutoff, $n = D$ [5]. Furthermore, in the limit of high spiking rates and a large number of neurons, we can model the dominant effect of the spike fluctuations using an equivalent noise level (like η above) [1]. However, as noted in [7], the behavior is very different if the neurons become synchronized.

3. A spiking model

In order to ensure that the above analysis holds for spiking networks, we can now model the network with leaky integrate-and-fire (LIF) neurons. For spiking neurons, we take the (instantaneous) firing rate, $a_i(t,\mathbf{A})$, to be the linearly filtered spike train produced by neuron i

$$a_i(t,\mathbf{A}) = \sum_n h(t - t_{i,n}(\mathbf{A})),$$

where the $t_{i,n}(\mathbf{A})$ are the spike times from neuron i indexed by n (up to t) given the function $f(x,\mathbf{A})$. These spike times are determined by the parameters of LIF neuron i, that define the non-linear encoding G_i. We take the linear filter, $h(t)$, to be the *PSC* (modeled as a simple RC circuit) produced by an incoming spike. Using the weights

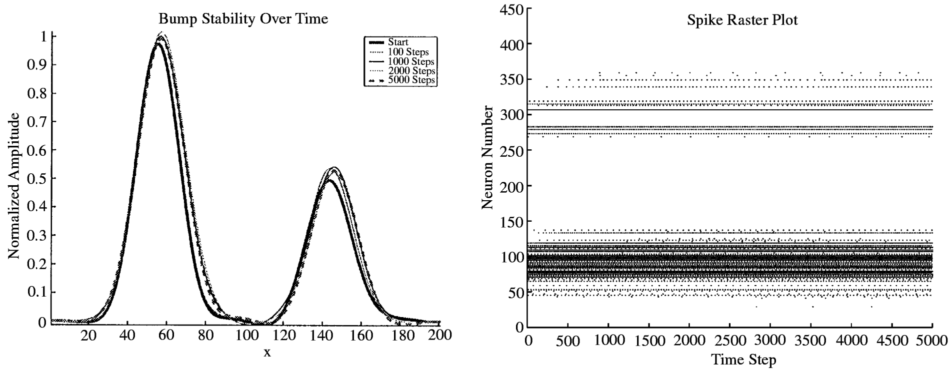

Fig. 1. The left panel shows the stability of two simultaneously encoded bumps of different amplitudes. The right panel shows the spike trains in the network encoding the bumps ($N = 200$; $\tau_{PSC} = 10$ ms).

we found early for the rate modeled LIF neurons, the current in the soma of neuron i can be written:

$$J_i(t, \mathbf{A}) = \alpha_i \sum_j \omega_{ij} a_j(t, \mathbf{A}) + \beta_i$$

$$= \alpha_i \sum_{n,j} \omega_{ij} h(t - t_{j,n}(\mathbf{A})) + \beta_i.$$

This current is used by the LIF neuron to determine spike the times, $t_{i,n}(\mathbf{A})$. The weights for this network are the same as for the rate model since the encoding and decoding procedures are the same in both cases. The results of a simulation based on this model is shown in Fig. 1. The tuning curves of the neurons used to construct this network are determined by randomly chosen LIF parameters, together with a randomly assigned, uniformly spaced Gaussian encoding function.

Lastly, we have simulated the analogous model in two dimensions. The analysis follows that outlined here, where Eqs. (1)–(4) can be re-written using vector notation. Not surprisingly, behavior of these networks is similar to that seen in the one-dimensional case.

4. Conclusion

We have shown how to generate networks of spiking neurons that can store a set of functions. We have given some examples of the kinds of analysis that can be carried out to help characterize the behavior of the neural network model. The particular set of functions we have employed consists of sets of Gaussian bumps of various amplitudes. These functions have a particularly interesting neurobiological interpretation. In areas that exhibit a kind of working memory, such as the lateral intraparietal area, there is evidence that multiple targets can be stored and that stimuli that differ along non-spatial dimensions (e.g. intensity or shape) give rise to different responses. Our model captures both of these aspects of this kind of working memory.

Interestingly, these networks have produced an experimental prediction. Currently, experimentalists tend to record only from neurons with stimuli at the center of their receptive fields. This results in a decrease in firing rate during the memory delay relative to the initial encoding of the stimulus. The networks we have generated reproduce this result, but they also include neurons whose firing rates increase when the stimuli is at the edge of the receptive field. We suspect, then, that a neuron tested with stimuli near the edges of its receptive field will increase its firing rate during the memory delay period.

References

[1] C.H. Anderson, Q.F. Huang, J. Clark, Harmonic analysis of spiking neuronal pairs, CNS ∗99, Pittsburgh, PA, 1999.

[2] A. Compte, N. Brunel, X.J. Wang, Spontaneous and spatially tuned persistent activity in a cortical working memory model, CNS ∗99, Pittsburg, PA, 1999.

[3] M. Camperi, X.J. Wang, A model of visuospatioal working memory in prefrontal cortex: recurrent network and cellular bistability, J. Comput. Neurosci. 5 (1998) 383–405.

[4] C. Eliasmith, C.H. Anderson, Developing and applying a toolkit from a general neurocomputational framework, Neurocomputing 26 (1999) 1013–1018.

[5] C. Eliasmith, C.H. Anderson, Rethinking central pattern generators: a general approach, Neurocomputing (in press).

[6] D. Hansel, H. Sompolinsky, Modeling feature selectivity in local cortical circuits, in: C. Koch, I. Segev (Eds.), Methods in Neuronal Modeling, 2nd Edition, MIT Press, Cambridge, MA, 1998.

[7] C.R. Laing, C.C. Chow, G.B. Ermentrout, Stationary bumps n networks of spiking neurons, Neural Comput., submitted for publication.

[8] M.L. Platt, G.W. Glimcher, Responses of intraparietal neurons to saccadic targets and visual distractors, J. Neurophysiol. 78 (1997) 1574–1589.

[9] A.B. Sereno, J.H.R. Maunsell, Shape selectivity in primate lateral intraparietal cortex, Nature 395 (1998) 500–503.

[10] L.H. Snyder, Personal communication.

[11] L.H. Snyder, A.P. Batista, R.A. Andersen, Coding of intention in the posterior parietal cortex, Nature 386 (1997) 167–170.

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