

Multi-packet regions in stabilized continuous attractor networks

Thomas P. Trappenberg and Dominic I. Standage

Faculty of Computer Science, Dalhousie University, Halifax, Canada

Abstract

Continuous attractor neural networks are recurrent networks with center-surround interaction profiles that are common ingredients in many neuroscientific models. We study realizations of multiple non-equidistant activity packets in this model. These states are not stable without further stabilizing mechanisms, but we show they can exist for long periods. While these states must be avoided in winner-take-all applications, they demonstrate that multiple working memories can be sustained in a model with global inhibition.

Key words: CANN, NMDA stabilization, bump networks, working memory

1 Introduction

Wilson and Cowan [1] derived a description of the population dynamics of neurons with excitatory and inhibitory pools coupled with center-surround interaction profiles. They identified various dynamical regimes in these networks and speculated that these regimes might map to different brain areas, including thalamic nuclei, visual neocortex, and prefrontal cortex. Continuous attractor neural networks (CANNs) are now common ingredients in models of information processing in the brain and are regarded as the principle model of cortical hypercolumns [2], place and head direction cells in the limbic system [3], and working memory [4].

The CANN model is most often used in a parameter regime where a single activity packet (also called bubble or bump) can be sustained without external input. The model then implements a winner-take-all function, which is appropriate for the modelling of place fields and feature representations in hypercolumns. Samsonovich [5] and Battaglia and Treves [6] have extended the model to multiple feature spaces, and our group has shown that many activity packets can be sustained simultaneously when the model is augmented with

biologically realistic stabilization mechanisms [7]. Multiple bubbles can also be sustained within a single topographic feature map with such stabilization mechanisms [8], which is relevant to the modelling of working memory. Here, we limit our study to a single feature space, and study the dependence of multiple activity packets on the amount of activity dependent global inhibition and the strength of the stabilization. We also show that multiple bubbles can be sustained for a considerable length of time without stabilization.

2 Methods

We consider a basic recurrent rate model with N nodes, though corresponding networks with spiking neurons have similar properties. The time evolution of the membrane state u_i of a node with index i is given by

$$\tau \frac{du_i(t)}{dt} = -u_i(t) + \sum_j w_{ij} r_j(t) \Delta x + I_i^{\text{ext}}(t), \quad (1)$$

where τ is a time constant, I_i^{ext} is the external input applied to the network, $\Delta x = 2\pi/N$ is a scale factor, and r_i is a rate that is related to u_i by sigmoidal gain function $g(u) = 1/(1 + \exp(-\beta(u - \alpha)))$ with a slope parameter $\beta = 0.1$ and firing threshold α . The weight matrix \mathbf{w} is determined in a learning phase with Hebbian learning, $w_{ij} \propto \sum_{\mu} r_i^{\mu} r_j^{\mu}$ on patterns with index μ . Such recurrent models are often studied after training on random patterns, resulting in networks with discrete attractors. In contrast, we study this model trained with well organized Gaussian patterns, where each pattern is centered around a different node in the network, $\mu = 1, \dots, N$. This results in an excitatory Gaussian weight matrix with width $\sigma_w = \sqrt{2}\sigma_r$,

$$r_i^{\mu} = A_r e^{-((i-\mu)\Delta x)^2/2\sigma_r^2} \rightarrow w_{ij}^{\text{ex}} = A_r \sqrt{\pi}\sigma_r e^{-((i-j)\Delta x)^2/4\sigma_r^2}, \quad (2)$$

which is then augmented by an inhibition constant C describing the activity dependent inhibition of an inhibitory pool of neurons and scaled by global strength constant A_w ,

$$w_{ij} = A_w \left(\frac{1}{A_r \sqrt{\pi}\sigma_r} w_{ij}^{\text{ex}} - C \right). \quad (3)$$

A well known problem in CANN models is that noise in the weight matrix leads to a drift of the activity packet [9]. It has been argued that drift slows down with increasing network sizes [3,4], and activity dependent bistabilities

in the excitability of neurons have also been shown to stabilize activity packets [10]. We implement the stabilization by a change of the threshold,

$$\Delta\alpha = \alpha_0\Theta(u), \quad (4)$$

where $\Theta(u)$ is the Heaviside function. This stabilization mechanism is sufficient to sustain neural activity after transient external stimuli without further (excitatory or inhibitory) support by other nodes in the neural layer [11]. A network without the lateral connections typical of CANN models, however, cannot implement the competition between stimuli that is essential to much of the brain processing for which CANNs were proposed. Here we follow our previous speculation that competition is the basis for the limited capacity of working memory [8], and study the dependence of the strength of the stabilization on multiple simultaneous activity packets in CANN models.

3 Results

First, we consider the model without stabilization. Appropriate values for C and A_w must be chosen to sustain an activity packet following transient input. This is illustrated in Figure 1A which shows the maximum node activity within the activity packet at time $t = 100\tau$ for various values of C and A_w . If inhibition is too weak then the entire network becomes active; too much inhibition shuts off all network activity. Additionally, the scaling parameter A_w must be strong enough to sustain an activity packet, but this constraint does not apply when a threshold activation function is used, as in the studies by Amari [12].

Amari’s analysis shows that a single activity packet is a stable attractor state in CANN networks [12], however, considerable time can be required before two or more externally driven activity regions merge or until one activity packet becomes dominant. The time required for a single maximum to emerge from two activated regions is shown in Figure 1B. The activity profile \mathbf{u} was measured at each iteration of dt/τ , where only two changes in direction of this curve determined a one-bubble state in the periodic network. The network was initialized with binary input bands around nodes 100 and 580, where one input band was 1% stronger than the other to brake the symmetry of input activity.

An example of 4 meta-stable asymmetric activity packets is shown in Figure 2A for a simulation with parameters $\sigma_r = 2\pi/80$, $A_w = 300$, and $C = 0.05$. No stabilizing bistability was implemented in this simulation. The results of a simulation with slightly increased inhibition constant, $C = 0.08$, are shown in

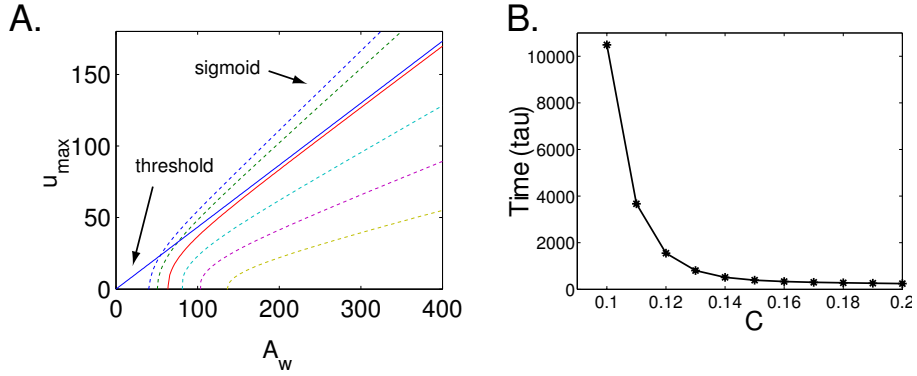


Fig. 1. (A) The maximal value of node activity u as a function of weight scaling factor A_w for different values of inhibition constant C . The curved solid line shows results with an inhibition constant of $C = 0.4$ for simulations with sigmoidal gain function $g(u) = (1 + \exp(-0.1u))^{-1}$. The straight solid line shows results from simulations with a threshold gain function $g(u) = \Theta(u)$ with otherwise unchanged parameters. Results from simulations with the sigmoidal gain function and different equidistant values for inhibition constant C are shown as dashed lines, from $C = 0.2$ (top) to $C = 0.7$ (bottom). (B) Time (τ) required for two activity packets to merge as a function of global inhibition (C) in the 1000 node network with input activity around nodes 100 and 580.

Figure 2B. Inhibition is now strong enough to introduce sufficient competition in the network such that one activity packet disappears within the time of $t = 100\tau$ following removal of the external stimulus.

We now consider the network with stabilization, and refer to the area where n equidistant activity packets do not considerably decay within time $t = 100\tau$ as an n -phase. The results of several simulations with varying values of a_0 and C are summarized in Figure 2C for different numbers of initial activity bands. Stabilizing more than 4 activity packets with the current parameters is impractical.

4 Discussion and conclusion

The results of the simulations shown in Figure 2 seem to contradict theoretical proof by Laing et al. [13,14] that without stabilizing mechanisms, more than one asymmetric activity packet cannot be sustained in the CANN model. However, our results simply show that multiple asymmetric bubbles can be sustained for considerable periods of time.

Figure 2C indicates that the effect of inhibition in the model is approximately inverse to that of stabilization. Increasing stabilization (increasing α_0) can lead to the stabilization of multiple activity packets, while increasing competition (increasing C) can destroy the existence of multiple activity packets.

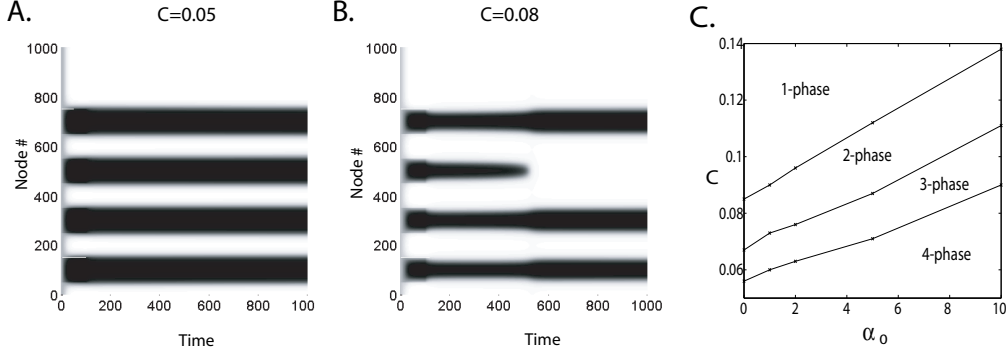


Fig. 2. (A) Network activity over time for a simulation with asymmetric external input at nodes 100, 300, 500, and 700 in a 1000 node network until $t = 100\tau$ with inhibition constant $C = 0.05$. (B) Corresponding simulations for $C = 0.08$. (C) Transitions between n -phases dependent on inhibition C and threshold adjustment α_0 . The transitions between the n and $n + 1$ phases were thereby studied with $n + 1$ of the original 4 input bands.

While this relationship appears linear, there are other parameters that effect stabilization in our model. Firstly, in the infinite time limit, we know there can be only one stable activity packet in the $\alpha_0 = 0$ limit. The relationship must therefore deviate considerably from the linear case for longer simulation times close to the α_0 limit. Secondly, the width of the weight profile and the strength of connectivity effects the number of activity packets sustained by the network, here modelled as $\sigma_r = 2\pi/80$ and $A_w = 200$ respectively. A wider weight profile and weaker connectivity serve to impede stabilization. Increasing the asymmetry of activation bands will further diminish stability of multi-packet solutions. Thirdly, the spatial discreteness inherent in numerical simulations has a stabilizing effect on the model. This effect increases with the sharpness of the activity packet profile, as only a Gaussian profile can be moved continuously with constant support over an equidistant lattice. This effect is directly related to strength parameter A_w , as increasing A_w saturates the activity packet profile.

While much recent attention has been paid to stabilizing activity packets in CANN models, little research has focused on balancing stabilization such that the model retains its winner-take-all functionality. Strong stabilization effectively partitions the network into a series of local networks with winner-take-all characteristics similar to networks with short range inhibition [15]. As such, it is possible to stabilize large numbers of activity packets under the CANN model. Alternatively, the use of low levels of global inhibition permits the co-existence of a small number of activity packets for finite periods in the absence of stabilization. We speculate that unlike models with short range inhibition, global inhibition in the CANN model parallels the effect of interaction between brain areas, limiting the number of simultaneous activity packets in accordance with the limited capacity of working memory.

References

- [1] H. Wilson, J. Cowan, A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue, *Kybernetik* 13 (1973) 55–80.
- [2] D. Hansel, H. Sompolsky, Modeling feature selectivity in local cortical circuits, 2nd Edition, MIT Press, 1998, Ch. 13, pp. 499 – 467.
- [3] K. Zhang, Representation of spatial orientation by the intrinsic dynamics of head-direction cell ensembles: a theory, *Journal of Neuroscience* 16 (4) (1996) 2112–2126.
- [4] A. Compte, N. Brunel, P. Goldman-Rakic, X.-J. Wang, Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model, *Cerebral Cortex* 10 (2000) 910–23.
- [5] A. Samsonovich, Attractor map theory of the hippocampal representation of space, Ph.D. thesis, University of Arizona (1997).
- [6] F. P. Battaglia, A. Treves, Attractor neural networks storing multiple space representations: A model for hippocampal place fields, *Physical Review* 58 (6).
- [7] S. Stringer, E. Rolls, T. Trappenberg, Self-organizing continuous attractor networks with multiple activity packets and the representation of space, *Neural Networks* 17 (2004) 5–27.
- [8] T. Trappenberg, Why is our capacity of working memory so large?, *Neural Information Processing: Letters & Reviews* 1(3) (2003) 97–101.
- [9] M. Tsodyks, T. Sejnowski, Associative memory and hippocampal place cells, *Int. Journal of Neural Systems* 6(supp. 1995) ((1995)) 81–86.
- [10] M. Camperi, X.-J. Wang, A model of visuospatial short-term memory in prefrontal cortex: cellular bistability and recurrent network, *J. Comput. Neurosci.* 5 (1998) 383–405.
- [11] J. Lisman, J. Fellous, X.-J. Wang, A role for nmda-receptor channels in working memory, *Nature Neurosci.* 1(4) (1998) 273–5.
- [12] S. Amari, Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological Cybernetics* 27 (1977) 77–87.
- [13] C. R. Laing, W. C. Troy, Two-bump solutions of amari-type models of neuronal pattern formation, *Physica D* 178.
- [14] B. G. Carlo R. Laing, William C. Troy, G. B. Ermentrout, Multiple bumps in a neuronal model of working memory, *SIAM J. Appl. Math.* 63 (1).
- [15] M. Usher, M. Stemmler, C. Koch, Z. Olami, Network amplification of local fluctuations causes high spike rate variability, fractal firing patterns and oscillatory local-field potentials, *Neural Computation* 6 (1994) 795–836.