

Robust persistent activity in neural fields with asymmetric connectivity

Cláudia Horta Wolfram Erlhagen

*Departamento de Matemática para C&T, Universidade do Minho, 4800-058
Guimarães, Portugal*

Abstract

Modeling studies have shown that recurrent interactions within neural networks are capable of self-sustaining non-uniform activity profiles. These patterns are thought to be the neural basis of working memory. However, the lack of robustness challenge this view as already small deviations from the assumed interaction symmetry destroy the attractor state. Here we analyze attractor states of a neural field model composed of bistable neurons. We show the existence of self-stabilized patterns that robustly represent the cue position in the presence of a substantial asymmetry in the connection profile. Using approximation techniques we derive an explicit expression for a threshold value describing the transition to a traveling activity wave.

Key words: Neural Field, Bistability, Working Memory, Spatial Orientation.

1 Introduction

Many models of stimulus-selective persistent activity in the brain are based on recurrent networks with 'continuous attractors'. The connectivity in these models supports the existence of self-stabilized activity profiles to represent any value along a continuous physical dimension such as direction or position [1,9,4]. A transient external input acts as a switch between a uniform rest state and one of the stable active states encoding a particular position or direction. Typically, the connections are organized in a Mexican-hat pattern with strong excitation between cells with similar preferred feature flanked by a strong "surround" inhibition. However, the continuity of the attractor states requires a perfect spatial symmetry in the connection profile (for review see [3]). Already

Email addresses: claudia@mct.uminho.pt (Cláudia Horta),
wolfram.erlhagen@mct.uminho.pt (Wolfram Erlhagen).

small deviations cause a drift in the spatial position of the activity pattern. To serve as a biologically plausible model for short term memory, the network must be sufficiently robust to imprecise components. On the other hand, it has been hypothesized in the context of head direction cells in rats that an asymmetry induced drift mechanism may be actively controlled by the biological system for an updating of spatial representations during self-motion. Vestibular or proprioceptive input, for instance, may have a modulatory effect on the synapses of the head direction cell network [9,?].

The main result of the present study is that the proposed model shows the required robustness but likewise allows implementing the dynamic updating mechanism. Concretely, our analytical and numerical results reveal: 1) the existence of a threshold for the asymmetry in the weight distribution below which static activity profiles exist, 2) that these activity patterns depend to some extent on the external input, and 3) that for larger asymmetries a transition to a traveling wave occurs.

2 Model description

We study a one-dimensional field model composed of bistable neurons with a non-symmetric, homogeneous connectivity of lateral inhibition type. As a concrete example we consider a field with periodic boundary conditions representing the circular space of heading direction. The time evolution of the network is governed by the equation:

$$\tau \frac{\partial u(x, t)}{\partial t} = f(u(x, t)) + \int_0^{360} w(x - y) u(y, t) dy + h, \quad (1)$$

where $\tau > 0$, $h < 0$ are constants defining the time scale and the resting level of the dynamics, respectively. The cubic like shaped non-linearity f describes the bistable behavior of each neuron with a stable resting and a stable excited state. To simplify the analysis we chose a piece-wise linear function given by (Fig. 1):

$$f(u) = \begin{cases} \frac{a}{2}u & \Leftarrow u \leq k/2 \\ -\frac{a}{2} \frac{k}{1-k} (u - 1/2) & \Leftarrow k/2 < u \leq 1 - k/2 \\ \frac{a}{2}(u - 1) & \Leftarrow u > 1 - k/2 \end{cases} \quad (2)$$

where $0 \leq k \leq 1$ and $a < 0$.

The synaptic weight distribution function w is chosen as a rectangular profile $w(y) = b_1(H(y + d_2) - H(y - d_1)) - b_2$, where H is the Heaviside step function ($H(x) = 1$ if $x > 0$ and is 0 otherwise) and $b_1 > 0$, $b_2 > 0$ are constants. The spatial range of the excitatory interactions to the “right” and to the “left” are described by $d_1 \geq 0$ and $d_2 \geq 0$, respectively, while the inhibition extends over the whole field. We assume $d_2 \leq d_1$ and set $A = b_1(d_1 + d_2)$ and $w_0 = A - b_2$. The *level of asymmetry* of the connection profile w is defined as the difference $d_1 - d_2$ relative to the total excitatory range $d_1 + d_2$, that is $ASY = \frac{d_1 - d_2}{d_1 + d_2}$, hence $0 \leq ASY \leq 1$. It is important to stress that the specific choice of the interaction profile w and the non-linearity f is motivated to allow for a more rigorous analysis of the steady state solutions. Numerical simulations show that smooth functions for f and w could have been chosen as well without qualitatively changing the results presented here.

3 Results

Because of the bistability of the neurons, a localized activity pattern triggered by a transient input appears to be discontinuous with a jump in the activation level occurring at positions represented by neurons x_1 and x_2 at the left and the right side, respectively (compare Fig. 2). We apply an approximation technique similar to [8] to derive explicit formulas for three parameters which describe the activity profile: the amplitude r , the width l , and the resting state u_b . As depicted in Fig. 2A, we use a piecewise linear function to approximate the neuronal pattern below and above the jump discontinuity. This allows us to directly evaluate the equilibrium solution of equation 1 for a neuron x_r at resting level, a fully excited neuron x_t and the two “transition neurons” x_1 and x_2 . We obtain the following set of equations:

$$\begin{cases} 0 = \frac{a}{2}u_b + w_0u_b - b_2lr + h \\ 0 = \frac{a}{2}u_1 + 2d_2b_1(u_1 - u_b) + b_1d_2 + w_0u_b - b_2lr + h \\ 0 = \frac{a}{2}u_b + \frac{a}{2}r - \frac{a}{2} + rA + w_0u_b - b_2lr + h \\ 0 = \frac{a}{2}u_2 + 2d_2b_1(u_2 - u_b) + b_1d_2 + (d_1 - d_2)r + w_0u_b - b_2lr + h. \end{cases}$$

Solving this system gives algebraic expressions for r , l and u_b ,

$$\begin{cases} r = \frac{a}{2A+a} \\ u_b = u_1 + \frac{b_1d_2}{2d_2b_1+a/2} \\ l = \frac{(a+2A)(w_0+a/2)}{ab_2}u_1 + \frac{a+2A}{-ab_2} \left(\frac{-(w_0+a/2)b_1d_2}{2d_2b_1+a/2} - h \right), \end{cases} \quad (3)$$

and also for the difference in the activation level of the two transition neurons, $u_1 = u(x_1)$ and $u_2 = u(x_2)$: $\Delta u \equiv u_2 - u_1 = \frac{-(d_1 - d_2)rb_1}{(2d_2b_1 + a/2)}$. Note that the specific choice of $d_1 = d_2$ represents the limit case of a symmetric weight distribution w and corresponds to $\Delta u = 0$. The dotted line in Fig. 2B illustrates for a concrete simulation example that the obtained values describe very well the static profile of the full model (solid line).

Two main properties of the neuronal patterns follow from previous expressions and from the observation that u_1 and u_2 must, because of the cubic shape of f , be in the range $[-k/2, k/2]$. First, we note that there is still a degree of freedom corresponding to the unspecified value of the left transition level u_1 in the expressions for l and u_b . Thus, each value of u_1 in the interval $[-k/2, k/2 - \Delta u]$ corresponds to a different localized equilibrium solution. While the height of the profile is completely determined by the network parameters, its width also depends, to some extent, on the external stimulation: using localized inputs with different spatial extent (e.g., gaussian functions, compare Fig. 3) allows a change in the width of the activity pattern between a minimum and a maximum value defined by the network parameters. Note that the interval of admissible widths of the profile decreases with increasing asymmetry (Fig. 4A, top).

Secondly, as illustrated in Fig. 3, static profiles do exist also in the presence of an asymmetric weight distribution w (dashed line). Shifting w to some extent to one side, that is choosing for instance $d_1 > d_2$, results in a non-uniform profile with an asymmetric shape. The condition $\Delta u \leq k$ for the existence of a static profile allows us to derive an analytic expression which describes the maximum level of asymmetry as a function of the network parameters:

$$ASY_h = \frac{k}{2A} \frac{(a + 2A)^2}{a(k - 1) + 2kA} \quad (4)$$

For $ASY > ASY_h$ the limit value for u_1 is reached and the profile starts to travel. As shown in Fig. 4A (bottom), the velocity of this form-stable traveling wave solution increases roughly linearly with ASY .

4 Discussion

In neural field models of lateral-inhibition type the cue position of a transient input is stored in the form of a continuum of network activity profiles. A major challenge we have addressed in the present work concerns the maintenance of activity in the presence of noise in the system. In a biological context, the memory function should work well even under the condition of damage or irregular training of the synaptic weights. We have shown that cellular

bistability in individual neurons drastically enhances robustness to imprecise components. In our modeling example, the integration of excitation from field neighbors at one side may exceed the summed excitation from the other side by as much as 56% of the total excitation. A similar impact (albeit smaller in absolute value) of cellular bistability on robustness in construction has been recently reported for networks encoding the cue by a rate code continuum [6]. Our results complement findings of previous simulation studies which reported that bistable neurons may render working memory robust against additive noise and distraction [4].

Reading out an asymmetric profile in the sense of a population code [5] results in a parameter value which does not represent the actual cue position. However, the bias in the direction of the asymmetry is significant only if the input appears to be extremely localized (Fig. 3B). For realistic population representations which cover an area of about 100 degrees or more (compare [9,4]) the bias can be neglected (Fig. 3A).

As suggested first by Zhang (1996), the self-stabilized traveling wave solution obtained with a sufficiently strong asymmetry may be used to update spatial representations during self-motion. A Hebbian learning scheme for establishing, during practice, the modulatory effect of additional information sources (e.g., vestibular or proprioceptive input) on the connection weights has been recently proposed by Stringer and colleagues [7].

In conclusion, the present framework offers new perspectives for applying and learning continuous attractor networks without biologically unrealistic constraints such as symmetric spatial interactions.

In future work we will generalize our analysis to two-dimensional fields. We expect that the gap in self-sustained activity profiles caused by the neuronal bistability will increase their robustness in a similar way as described here. From the viewpoint of a realistic cortical modelling, another important extension is to relax also the assumption about the homogeneity of the spatial interactions. It is known that even a weak heterogeneity may affect the existence and stability of patterns [2].

Acknowledgements

This work was supported by the European grant “JAST” (IST-2-003747-IP).

References

- [1] S. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, (27):77–87, 1977.

- [2] P.C. Bressloff. Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Physica D*, 155:83–100, 2001.
- [3] C. D. Brody, R. Romo, and A. Kepecs. Basic mechanisms for graded persistent activity: Discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology*, 13:204–211, 2003.
- [4] M. Camperi. and X.-J. Wang. A model of visuospatial working memory in prefrontal cortex: Recurrent network and cellular bistability. *Journal of Computational Neuroscience*, (5):383 – 405, 1998.
- [5] W. Erlhagen, A. Bastian, D. Jancke, A. Riehle, and G. Schöner. The distribution of neuronal population activation as a tool to study interaction and integration in cortical representations. *Journal of Neuroscience Methods*, 94:53–66, 1999.
- [6] A. A. Koulakov, S. Raghavachari, A. Kepecs, and J. E. Lisman. Model for a robust neural integrator. *Nature Neuroscience*, 5(8):775–782, 2002.
- [7] S. M. Stringer, T. P. Trappenberg, E. T. Rolls, and I. E. T. de Araújo. Self-organizing continuous attractor networks and path integration: one-dimensional models of head direction cells. *Network: Computation in Neural Systems*, 13:217–242, April 2002.
- [8] T. Wennekers. Orientation tuning properties of simple cells in area v1 derived from an approximative analysis of nonlinear neural field models. *Neural Computation*, 13:1721–1747, 2001.
- [9] K. Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *The Journal of Neuroscience*, 16(6):2112–2126, March 1996.

Biosketch:

Cludia Horta received her MSc. in Computer Science in 1999 from the University of Minho (UM) in Portugal. She is currently finishing a Ph.D. project at the Department of Mathematics for Science and Technology at UM in the area of non linear dynamical systems. Her research interests include the analysis of pattern formation in coupled systems of integro-differential equations with applications both in Neuroscience and Robotics.

Wolfram Erlhagen is an Associate Professor at the Department of Mathematics for Science and Technology and director of the research centre *Officina Mathematica* at the University of Minho in Portugal. He studied Mathematics and Physics at the Universities of Bonn and Bochum in Germany, obtaining a doctoral degree in Mathematics in 1996 with a work on pattern formation in active media. Before coming to Portugal he was leader of the Theoretical Biology group at the Institute for Neuroinformatics, Ruhr-University Bochum (1996-2000) and invited researcher at the CNRS in Marseille, France (1997).

Dr Erlhagen's research interests cover the fields of spatially extended dynamical systems, neurocomputational and cognitive modelling, and architectures for autonomous robots. In close cooperation with experimental groups he applies his theoretical investigations to problems of motor planning, visual perception and reasoning.

Figure captions:

Figure 1: Sketch of the non-linear function $f(u)$ for different values of k .

Figure 2: A) Piecewise linear approximation (solid line) for the activity profile below and above the jump discontinuity (dashed line). B) Plot of a self-sustained activity profile in response to a transient input centered at $x = 0$ (solid line). The dashed line indicates the values for l , r and u_b obtained analytically by solving system 3.

Figure 3: Stable activity profiles in response to transient gaussian inputs (dashed-dotted lines) with different standard deviation σ centered at $x = 0$ for a symmetric (solid line) and an asymmetric field connectivity (dashed line).

Figure 4: A) On top, the maximum (dotted line) and the minimum (solid line) possible width of the profile are shown as a function of the asymmetry ASY . At $ASY = 0.56$ the transition to a traveling wave occurs the speed of which increases roughly linearly with increasing asymmetry (bottom). B) Evolution of a traveling wave in response to a transient input ($ASY = 0.6$).







