Models of networks - continued

Readings: D&A, chapter 7.

Network models - summary

- Network models: to understand the implications of connectivity in terms of computation and dynamics.
- 2 Main strategies: Spiking vs Firing rate models.
- The issue of the emergence of orientation selectivity as a model problem, extensively studied theoretically and experimentally.
  - Two main models: feed-forward and recurrent.
  - Detailed spiking models have been constructed which can be directly compared to electrophysiology
  - The same problem is also investigated with a firing rate model, a.k.a. the 'ring model'.

Thursday, 4 February 2010

The Ring Model (1)

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 3844–3848, April 1995
Neurobiology

Theory of orientation tuning in visual cortex

(neural networks/cross-correlations/symmetry breaking)

R. BEN-YISHAI*, R. LEV BAR-OR*, AND H. SOMPOLINSKY†

*Racah Institute of Physics and Center for Neural Computation, Hebrew University, Jerusalem 91904, Israel; and †AT&T Bell Laboratories, Murray Hill, NJ 07974

Communicated by Pierre C. Hohenberg, AT&T Bell Laboratories, Murray Hill, NJ, December 21, 1994 (received for review July 28, 1994)

ABSTRACT **The role of intrinsic cortical connections in processing sensory input and in generating behavioral output is poorly understood. We have examined this issue in the context of the tuning of neuronal responses in cortex to the orientation of a visual stimulus. We analytically study a simple network model that incorporates both orientation-selective input from the lateral geniculate nucleus and orientation-specific cortical interactions. Depending on the model parameters, the network exhibits orientation selectivity that originates from within the cortex, by a symmetry-breaking mechanism. In this case, the width of the orientation tuning can be sharp even if the lateral geniculate nucleus inputs are only weakly anisotropic. By using our model, several experimental consequences of this cortical mechanism of orientation tuning are derived. The tuning width is relatively independent of the contrast and angular anisotropy of the visual stimulus. The transient population response to changing of the stimulus orientation exhibits a slow "virtual rotation." Neuronal cross-correlations exhibit long time tails, the sign of which depends on the preferred**

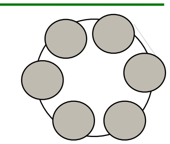
ivity among cortical neurons can be gained from measurements of the correlations between the responses of different neurons (10). Theoretical predictions regarding the magnitude and form of correlation functions in neuronal networks have been lacking.

Here we study mechanisms for orientation selectivity by using a simple neural network model that captures the gross architecture of primary visual cortex. By assuming simplified neuronal stochastic dynamics, the network properties have been solved analytically, thereby providing a useful framework for the study of the roles of the input and the intrinsic connections in the formation of orientation tuning in the cortex. Furthermore, by using a recently developed theory of neuronal correlation functions in large stochastic networks, we have calculated the cross-correlations (CCs) between the neurons in the network. We show that different models of orientation selectivity may give rise to qualitatively different spatiotemporal patterns of neuronal correlations. These predictions can be tested experimentally.

Model

Thursday, 4 February 2010

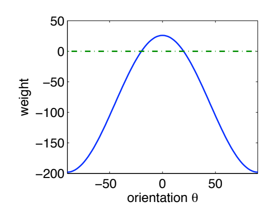
The Ring Model (2)

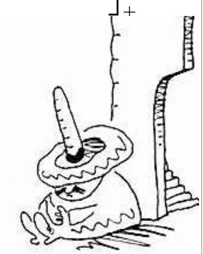


- N neurons, with preferred angle, θ_i ,evenly distributed between $-\pi/2$ and $\pi/2$
- Neurons receive thalamic inputs *h*.

+ recurrent connections, with excitatory weights between nearby cells and inhibitory weights between cells that are further apart (mexican-hat profile)

$$\tau_r \frac{dv(\theta)}{dt} = -v(\theta) + \left[h(\theta) + \int_{-\pi/2}^{\pi/2} \frac{d\theta'}{\pi} \left(-\lambda_0 + \lambda_1 \cos(2(\theta - \theta')) \right) v(\theta') \right]_+$$



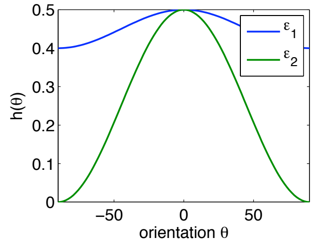


Thursday, 4 February 2010

The Ring Model (3)

- h is input, can be tuned (Hubel Wiesel scenario) or very broadly tuned.

$$h(\theta) = c[1 - \epsilon + \epsilon * cos(2\theta)]$$



- The steady-state can be solved analytically.

Model analyzed like a physical system.

- Model achieves i) orientation selectivity; ii) contrast invariance of tuning, even if input is very broad.
- The width of orientation selectivity depends on the shape of the mexican-hat, but is independent of the width of the input.
- Symmetry breaking /Attractor dynamics.

The Ring Model (4)

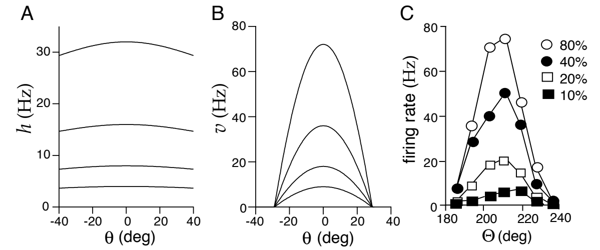
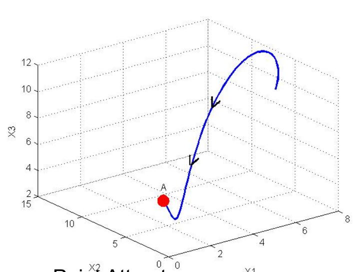


Figure 7.10: The effect of contrast on orientation tuning. A) The feedforward input as a function of preferred orientation. The four curves, from top to bottom, correspond to contrasts of 80%, 40%, 20%, and 10%. B) The output firing rates in response to different levels of contrast as a function of orientation preference. These are also the response tuning curves of a single neuron with preferred orientation zero. As in A, the four curves, from top to bottom, correspond to contrasts of 80%, 40%, 20%, and 10%. The recurrent model had $\lambda_0 = 7.3$, $\lambda_1 = 11$, $A = 40$ Hz, and $\epsilon = 0.1$. C) Tuning curves measure experimentally at four contrast levels as indicated in the legend. (C adapted from Sompolinsky and Shapley, 1997; based on data from Sclar and Freeman, 1982.)

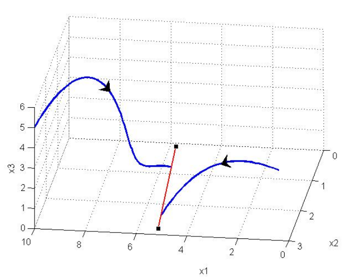
Attractor Networks

- Attractor network : a network of neurons, usually recurrently connected, whose time dynamics settle to a stable pattern.
- That pattern may be stationary (fixed points), time-varying (e.g. cyclic), or even stochastic-looking (e.g., chaotic).
- The particular pattern a network settles to is called its 'attractor'.

- The ring model is called a line (or ring) attractor network. Its stable states are also sometimes referred to as 'bump attractors'.



Point Attractor



Line Attractor

Pergamon

PII: S0042-6989(97)00100-4

Vision Res., Vol. 37, No. 21, pp. 3061–3071, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Predictions of a Recurrent Model of Orientation Selectivity

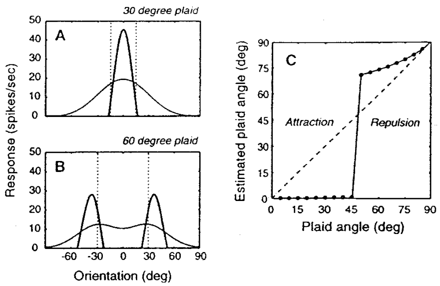
MATTEO CARANDINI,*† DARIO L. RINGACH‡

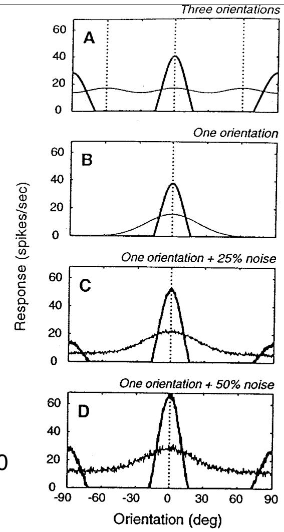
Received 3 January 1997; in revised form 21 March 1997

Recurrent models of orientation selectivity in the visual cortex postulate that an initially broad tuning given by the pattern of geniculate afferents is substantially sharpened by intracortical feedback. We show that these models can be tested on the basis of their predicted responses to certain visual stimuli, without the need for pharmacological or physiological manipulations. First, we consider a detailed recurrent model proposed by Somers, Nelson and Sur [(1995) *Journal of Neuroscience, 15,* 5448–5465] and show that it can be simplified to a single equation: a center–surround feedback filter in the orientation domain. Then, we explore the responses of the simplified model to stimuli containing two or more orientations. We find that the model exhibits peculiar responses to stimuli containing two orientations, such as plaids or crosses: if the component orientations differ by less than 45 deg the model cannot distinguish between them; if the orientations differ by more than 45 deg the model overestimates their angle by as much as 30 deg. Moreover, the model cannot signal the presence of three orientations separated by 60 deg (it responds as if there were only two orientations), and the addition of two-dimensional visual noise to an oriented stimulus results in strong spurious responses at the orthogonal orientation. We argue that the effects of attraction and repulsion between orientations and the emergence of responses at off-optimal orientations are common to a wide class of feedback models of orientation selectivity. These models could thus be tested by measuring the visual responses of cortical neurons to stimuli containing multiple orientations. © 1997 Elsevier Science Ltd

Orientation Striate cortex Model Plaid Noise

- Reduction of the spiking model of Somers et al 1995 to rate model



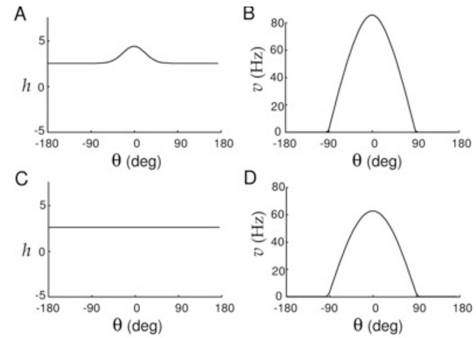


- Model was tested with stimuli containing more than 1 orientation (crosses)
- Model fails to distinguish angles separated by 30 deg, overestimates larger angles
- spurious attractors with noise

matlab code available online : http://www.carandinilab.net/publications

The Ring Model (5): Sustained Activity

- If recurrent connections are strong enough, the pattern of population activity once established can become independent of the structure of the input. It can persists when input is removed.
- A model of working memory ?



Sustained activity, Working Memory, Associative memory

Readings:

C.Constandinis and XJ Wang, , "a neural circuit basis for spatial working memory", Neuroscientist, 2004

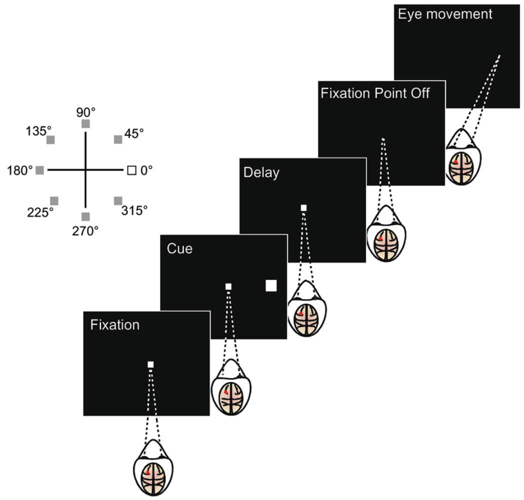
What is working memory ? (a.k.a. short-term memory)

- The ability to hold information over a time scale of seconds to minutes
- a critical component of cognitive functions (language, thoughts, planning etc..)

Delayed match-to sample task:
remember 'red'

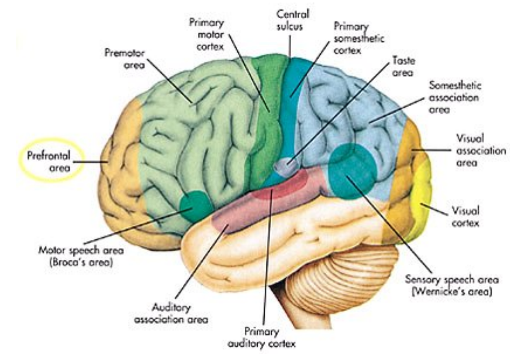
Oculo-motor delayed response task:

remember location of cue.

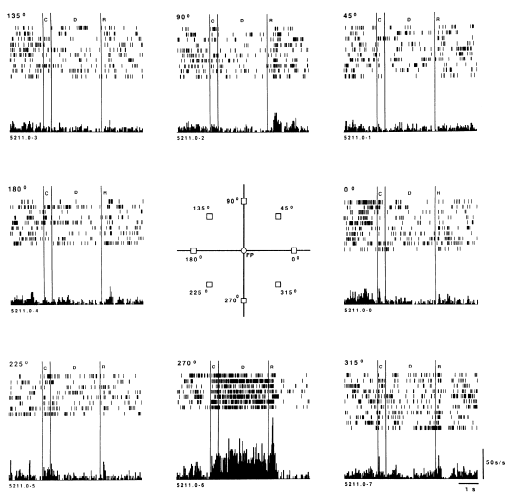


Sustained activity in PFC (1)

• Lesion and inactivation studies demonstrate crucial role of Prefrontal Cortex (PFC) in working memory, in particular dorsolateral PFC (PFdl).



Sustained activity in PFC (2)



Funahashi et al, 1989

Working memory vs Long-term memory

• Long-term memory : molecular or structural changes

• Short-term/ working memory: dynamic process that has not yielded to molecular characterization. Sustained Activity.

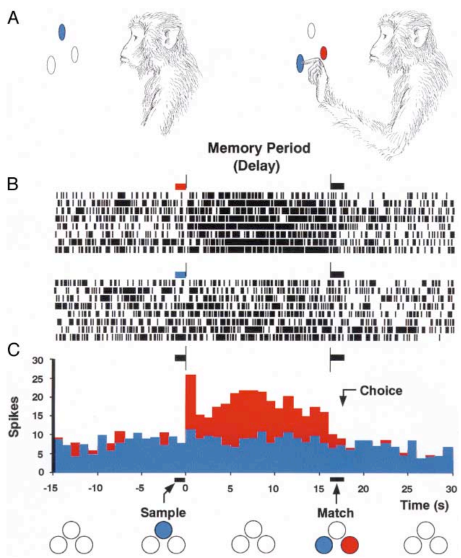
Working memory vs Long-term memory

- Long-term memory : molecular or structural changes
- Short-term/ working memory: dynamic process that has not yielded to molecular characterization. Sustained Activity.

Sustained activity is very widespread

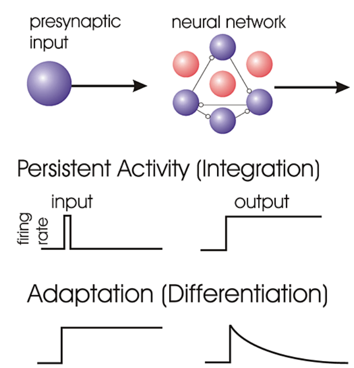
- Sustained activity is a widespread phenomenon
- LIP and PP also have neurons which direction-specific memory fields, similar to PFC.
- Also found in inferotemporal cortex (IT), see e.g. Fuster and Jervey 1982.
  Example of a discrete working memory.
- Memory related activity is also described in V3A, MT, V1, entorhinal cortex, Pre motor cortex, SMA, SC, basal ganglia...
- The distinct and cooperative roles of these areas remain unresolved.

Sustained activity in IT



Fuster and Jervey 1982

Brain calculus : integration and differentiation



- integration (persistent activity) seems to be mainly due to network interactions, while differentiation (adaptation) seems mainly cellular and synaptic depression

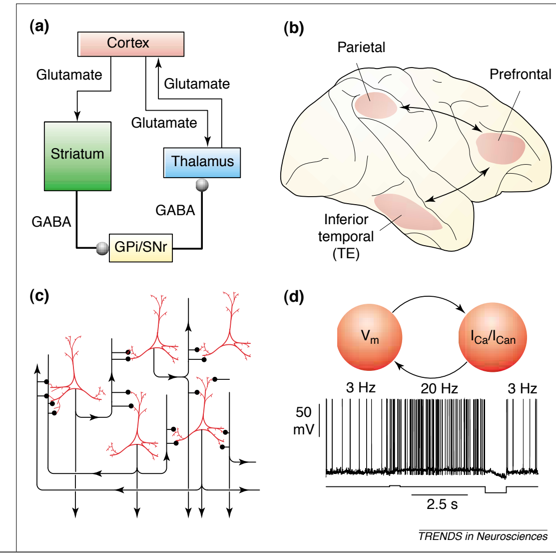
Working Memory and Sustained Activity

- A theory of working memory should answer:
- how it is initiated?
- why does it persist ?
- what makes it specific?
- how does it end?

- reason for capacity limit?
- relationship with attention, long term memory?

- Mechanism : reverberations through connections (which?), or cellular?
- Lots of experimental and theoretical work to answer these questions, in PFC, HD, Oculo-motor system

How does a transient stimulus cause a lasting change in neural activity?



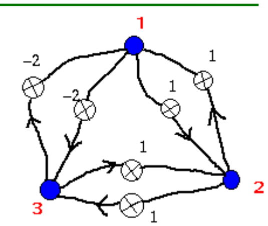
Attractor paradigm for persistent activity

- Since the 1970s it has been proposed that delay activity patterns can be theoretically described by 'dynamical attractors'

Hopfield Networks

- A Hopfield net is a form of recurrent artificial neural network invented by John Hopfield (1982).
- Hopfield nets typically have binary (1/-1 or 1/0) threshold units:

$$s_i = \begin{cases} 1 & \text{if } \sum_j w_{ij} s_j > \theta_i, \\ -1 & \text{otherwise.} \end{cases}$$



weights in black
Nodes numbers in red

where s_j state of unit j, and θ_i is the threshold

The weights have to follow: $w_{ii}=0$, $w_{ij}=w_{ji}$

- Hopfield nets have a scalar value associated with each state of the network referred to as the "energy", E, of the network, where:

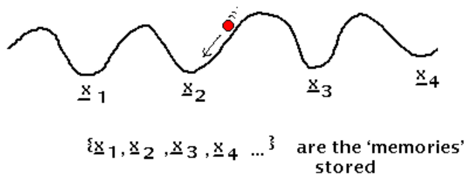
$$E = -\frac{1}{2}\sum_{i<j} w_{ij} s_i s_j + \sum_i \theta_i \; s_i$$

Hopfield Networks

• Running: at each step, pick a node at random and update (asynchronous update)
The energy is guaranteed to go down and the network to settle in local minima of the energy function.

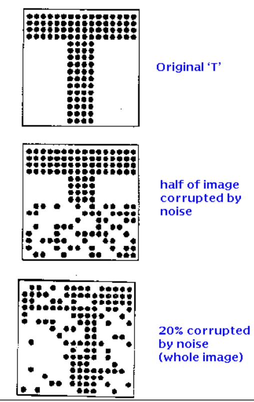
• Learning: the weights are learnt, so as to 'shape' those local minima.
The network will learnt to converge to learnt state even if it is given only part of the state

$$w_{ij} = \frac{1}{N} \sum_{k=1}^{k=N} \xi_i^k \xi_j^k$$



Associative memories

• The Hopfield network is an associative/content addressable memory. It can be used to recover from a distorted input the trained state that is most similar to that input. E.g., if we train a Hopfield net with 5 units so that the state (1, 0, 1, 0, 1) is an energy minimum, and we give the network the state (1, 0, 0, 0, 1) it will converge to (1, 0, 1, 0, 1).



Attractor paradigm for persistent activity

• Since the 1970s it has been proposed that delay activity patterns can be theoretically described by 'dynamical attractors'

• Recently, a great effort to build biophysically plausible model of sustained activity / attractor dynamics for memory.

The Ring Model (1)

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 3844–3848, April 1995
Neurobiology

Theory of orientation tuning in visual cortex

(neural networks/cross-correlations/symmetry breaking)

R. Ben-Yishai*, R. Lev Bar-Or*, and H. Sompolinsky†

*Racah Institute of Physics and Center for Neural Computation, Hebrew University, Jerusalem 91904, Israel; and †AT&T Bell Laboratories, Murray Hill, NJ 07974

Communicated by Pierre C. Hohenberg, AT&T Bell Laboratories, Murray Hill, NJ, December 21, 1994 (received for review July 28, 1994)

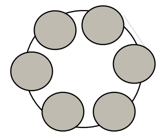
ABSTRACT The role of intrinsic cortical connections in processing sensory input and in generating behavioral output is poorly understood. We have examined this issue in the context of the tuning of neuronal responses in cortex to the orientation of a visual stimulus. We analytically study a simple network model that incorporates both orientation-selective input from the lateral geniculate nucleus and orientation-specific cortical interactions. Depending on the model parameters, the network exhibits orientation selectivity that originates from within the cortex, by a symmetry-breaking mechanism. In this case, the width of the orientation tuning can be sharp even if the lateral geniculate nucleus inputs are only weakly anisotropic. By using our model, several experimental consequences of this cortical mechanism of orientation tuning are derived. The tuning width is relatively independent of the contrast and angular anisotropy of the visual stimulus. The transient population response to changing of the stimulus orientation exhibits a slow "virtual rotation." Neuronal cross-correlations exhibit long time tails, the sign of which depends on the preferred

ivity among cortical neurons can be gained from measurements of the correlations between the responses of different neurons (10). Theoretical predictions regarding the magnitude and form of correlation functions in neuronal networks have been lacking.

Here we study mechanisms for orientation selectivity by using a simple neural network model that captures the gross architecture of primary visual cortex. By assuming simplified neuronal stochastic dynamics, the network properties have been solved analytically, thereby providing a useful framework for the study of the roles of the input and the intrinsic connections in the formation of orientation tuning in the cortex. Furthermore, by using a recently developed theory of neuronal correlation functions in large stochastic networks, we have calculated the cross-correlations (CCs) between the neurons in the network. We show that different models of orientation selectivity may give rise to qualitatively different spatiotemporal patterns of neuronal correlations. These predictions can be tested experimentally.

Model

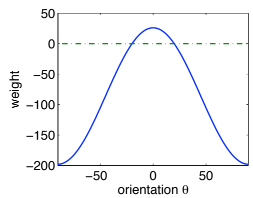
The Ring Model (2)

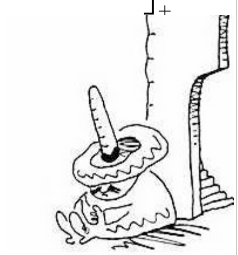


- N neurons, with preferred angle, θ_i ,evenly distributed between $-\pi/2$ and $\pi/2$
- Neurons receive thalamic inputs *h.*

+ recurrent connections, with excitatory weights between nearby cells and inhibitory weights between cells that are further apart (mexican-hat profile)

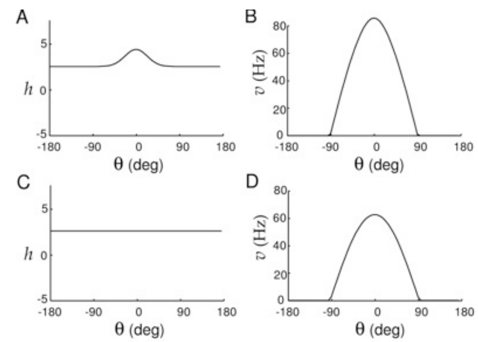
$$\tau_r \frac{dv(\theta)}{dt} = -v(\theta) + \left[h(\theta) + \int_{-\pi/2}^{\pi/2} \frac{d\theta'}{\pi} \left(-\lambda_0 + \lambda_1 \cos(2(\theta - \theta'))\right) v(\theta') \right]_+$$





The Ring Model (5): Sustained Activity

- If recurrent connections are strong enough, the pattern of population activity once established can become independent of the structure of the input. It can persists when input is removed.
- A model of working memory ?



Network Mechanisms & Biophysical Models

- Anatomical organization of PFC resembles a recurrent network
- Biophysical realistic computational modeling has shown that such recurrent networks can give rise to location-specific, persistent discharges (Compte et al 2000, Gutkin et al 2000, Tegner et al 2002, Renart et al 2003a, Wang et al 2004)

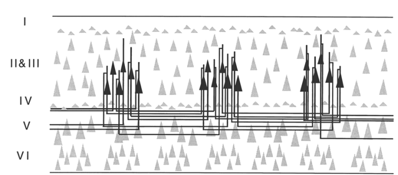


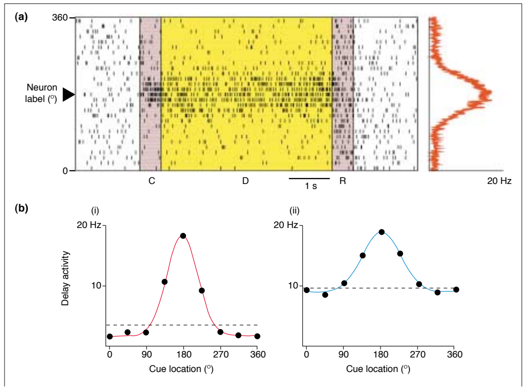
Fig. 4. Schematic diagram illustrating the pattern of connections between prefrontal neurons in the superficial layers. The figure summarizes results of anatomical tracer injection experiments and retrograde labeling. From Kritzer and Goldman-Rakic (1995), with permission.

Network Mechanisms & Biophysical Models

- Modeling studies show that stability is an issue in such network.
- Strong recurrent inhibition is needed to prevent runaway excitation and maintain specificity
- Models are also challenged by accounting for spontaneous activity in addition to memory state
- Oscillations can destabilize the memory activity.

- Working memory is found to be particularly stable when excitatory reverberations are characterized by a fairly slow time course, e.g. when synaptic transmission is mediated by NMDA receptors (prediction)

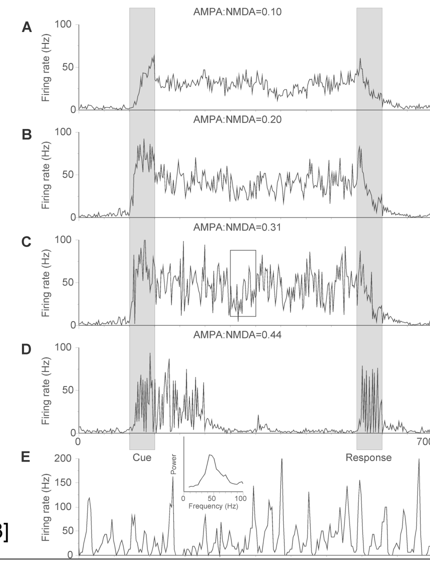
Network Mechanisms & Biophysical Models



[Compte, Brunel, Goldman-Rakic and Wang, 2000]
Network of ~2500 integrate and fire neurons, mexican hat connectivity, NMDA excitation.

Network Mechanisms & Biophysical Models

Fig. 6. Stability of persistent activity as a function of the AMPA:NMDA ratio at the recurrent excitatory synapses. *A–D*, Temporal course of the average firing rate across a subpopulation of cells selective to the presented transient input, for different levels of the AMPA:NMDA ratio. As the ratio is increased, oscillations of a progressively larger amplitude develop during the delay period, which eventually destabilize the persistent activity state. *E*, Snapshot of the activity of the network in (*C*) between 3 and 3.5 seconds. *Top*, Average network activity. *Bottom*, Intracellular voltage trace of a single neuron. *Inset*, Power spectrum of the average activity of the network, showing a peak in the γ (40 Hz) frequency range. Persistent activity is stable even in the presence of synchronous oscillations. Adapted with permission from Renart, Brunel, and others (2003).



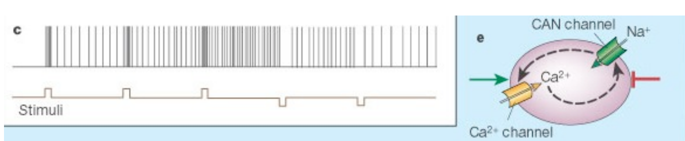
[Renart, Brunel, Wang , 2003]

But cellular mechanisms should not be forgotten ...

[Egorov et al, Nature, 2002]

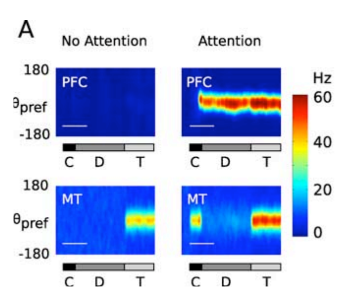
- Layer 5 of EC in vitro, intracellular depolarization + bath application of the ACh-receptor agonist leads to a Ca2+ -dependent plateau potential.
- This leads to sustained firing at a constant rate > 13 min
- independent of synaptic transmission.
- Level of activity can be increased or decreased using repeated inputs.

Could attractors be suited for remembering learned stimuli while such a system could help maintaining new stimuli?



Lots of interesting questions

- How are these attractors learnt?
- What is the relation with Attention?
- What is the relation with Long-term Memory ? (Is sustained activity helpful for storage of memory?)



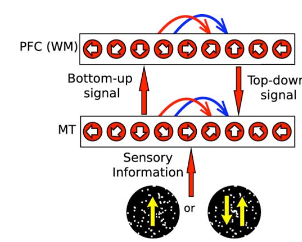
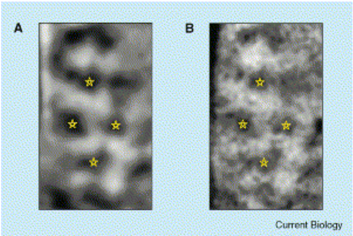


Figure 1. Scheme of the loop architecture (red is excitation, and blue is inhibition). Two kinds of motion stimuli are considered (random-dot patterns; yellow arrows indicate signal

Ardid, Wang and Compte 2007

A related problem: spontaneous activity

- Where does it come from?
- How is it maintained? How does it 'move'?
- Are these 'attractor states'?
- Is it structured?
- Why is it there? (any functional advantages?)
- Is it noise?
- Is it the brain trying to 'predict' the input?



evoked (horizontal orientation) | spontaneous (one frame)

Arieli et al 1997; Tsodyks et al, 1999;
Fiser et al, Nature, 2004