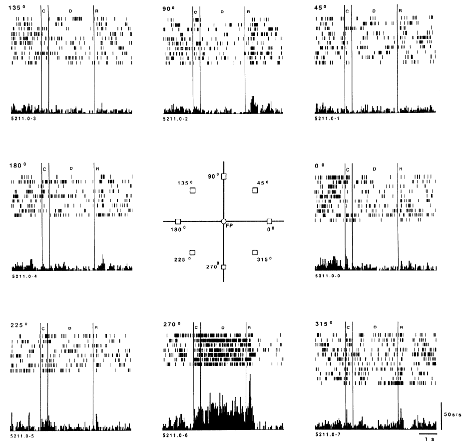
Sustained activity, Working Memory, Associative memory

Readings:

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

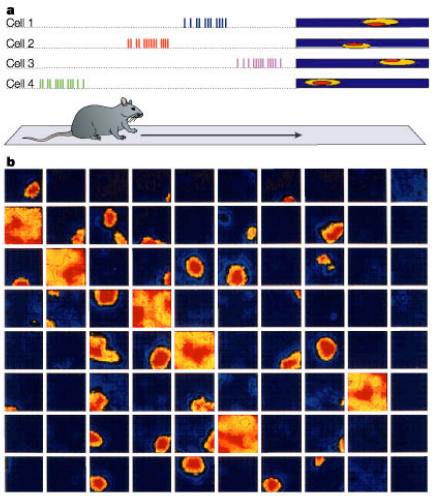
Sustained activity in PFC

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



[Funahashi et al, 1989]

Place cells in hippocampus



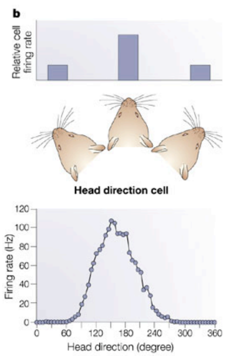
Nature Reviews | Neuroscience

Place cells are principal neurons in the hippocampus that fire strongly whenever an animal is in a specific location in an environment corresponding to the cell's "place field".

• often direction-selective
• suggests that the primary function of the rat hippocampus is to form a cognitive map of the rat's environment

• visual cues seem to be the primary determinant of place cell firing, but firing persists in the dark, suggesting that proprioception or other senses contribute as well.

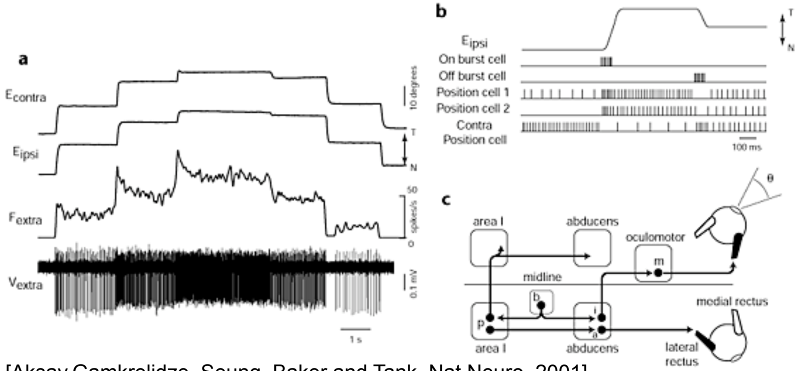
Head-direction cells



Neurons that are active only when the animal's head points in a specific direction within an environment.
These cells are found in many different structure of the limbic system.
Also continue to fire in darkness.

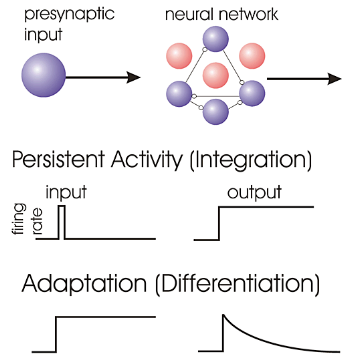
Neural integrator in the Oculomotor System

• in a premotor area that is responsible for holding the eyes still during fixation, persistent neural firing encodes the angular position of the eyes in a characteristic fashion: below a threshold position the neural is silent, and above it, the firing rate is linearly related to position.



[Aksay,Gamkrelidze, Seung, Baker and Tank, Nat Neuro, 2001]

Brain calculus : integration and differentiation



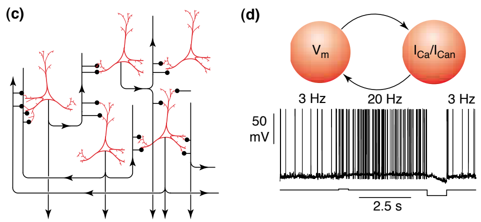
• while integration (persistent activity)seems to be mainly due to network interactions, differentiation (adaptation) seems mainly cellular and synaptic depression

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 ends?

- 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, Oculomotor system



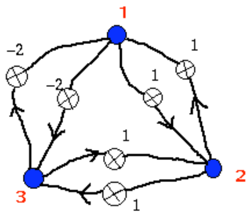
Attractor paradigm for persistent activity

• 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:

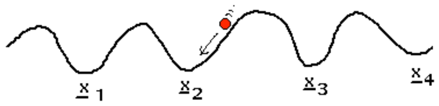
$$E = -\frac{1}{2}\sum_{i<j} w_{ij} s_i s_j + \sum_i \theta_i\ s_i$$

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

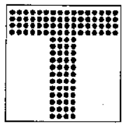
$$w_{ij} = \frac{1}{N}\sum_{k=1}^{k=N} \xi_i^k \xi_j^k$$



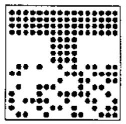
$\{\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4 \ldots\}$ are the 'memories' stored

Associative memories

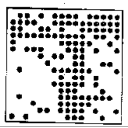
• The Hopfield network is an associative/content addressable memory. It can be used to recover from a distorted input the trained state that is most similar to that input. E.g., if we train a Hopfield net with 5 units so that the state (1, 0, 1, 0, 1) is an energy minimum, and we give the network the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



Original 'T'



half of image corrupted by noise



20% corrupted by noise (whole image)

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.

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)

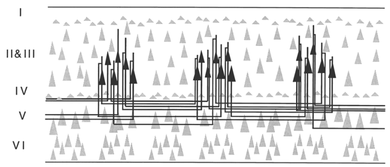


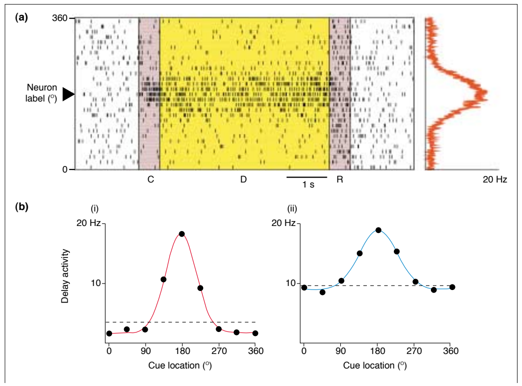
Fig. 4. Schematic diagram illustrating the pattern of connections between prefrontal neurons in the superficial layers. The figure summarizes results of anatomical tracer injection experiments and retrograde labeling. From Kritzer and Goldman-Rakic (1995), with permission.

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 reverberation are characterized by a fairly slow time course, e.g. when synaptic transmission is mediated by NMDA receptors (prediction)

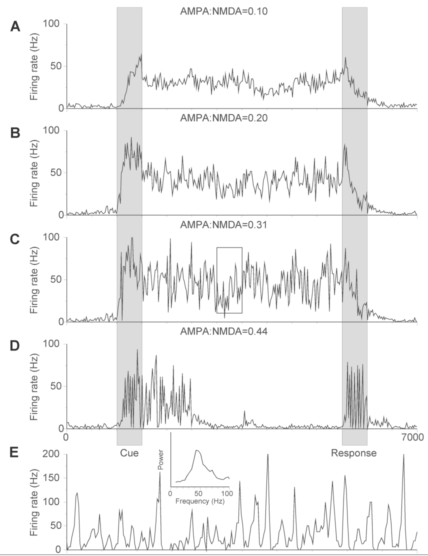
Network Mechanisms & Biophysical Models



[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]

Network Mechanisms & Biophysical Models

Box 2. Outstanding questions

Recent theoretical models have raised several neurophysiological questions that can be investigated experimentally. Answers to these questions will help to elucidate the mechanisms of neural persistent activity.

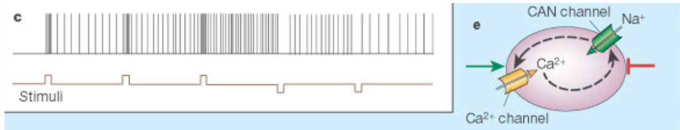
- **What is the minimum anatomical substrate of a reverberatory circuit capable of persistent neural activity?**
- **Is persistent activity primarily sustained by synaptic reverberation, or by bistable dynamics of single neurons?**
- **What is the NMDA:AMPA ratio at recurrent synapses of association cortices, especially in the prefrontal cortex?**
- **How does this ratio depend on the frequency of repetitive stimulation and on neuromodulation?**
- **What are the negative feedback mechanisms responsible for the rate control in a working memory network?**
- **Is delay period activity asynchronous between neurons, or does it display partial network synchrony and coherent oscillations?**
- **Is delay period activity more sensitive to NMDAR antagonists compared with AMPAR antagonists?**
- **Does persistent activity disappear in an abrupt fashion, with a graded block of NMDAR and AMPAR channels, as predicted by the attractor model?**
- **How significant are drifts of persistent activity during working memory? Are drifts random or systematic over trials?**
- **What are the biological mechanisms underlying the robustness of a memory network with a continuum of persistent activity patterns?**

But cellular mechanisms should not be forgotten ...

[Egorov et al, Nature, 2002]

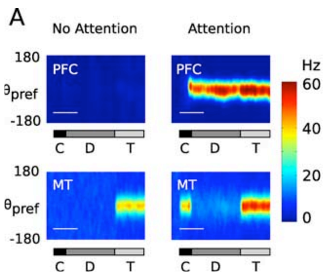
- Layer 5 of EC in vitro, intracellular depolarization + bath application of the ACh-receptor agonist leads to a Ca2+ -dependent plateau potential.
- This leads to sustained firing at a constant rate > 13 min
- independent of synaptic transmission.
- Level of activity can be increased or decreased using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

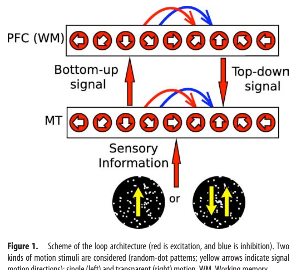
Could attractors be suited for remembering learned stimuli while such a system could help maintaining new stimuli?



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?)

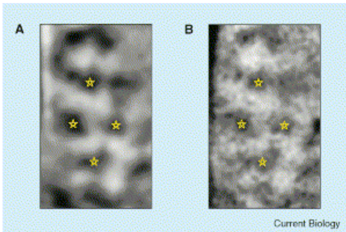




Ardid, Wang and Compte 2007

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