Short-term facilitation stabilizes the working memory trace.

Vladimir Itskov[1], David Hansel[2], Misha Tsodyks[3]

[1]Columbia University; [2]CNRS-Université Paris Descartes and Hebrew University of Jerusalem; [3]Weizmann Institute of Science.

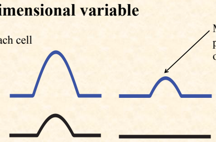
Working memory for a continuous variable.
Persistent activity observed in cortex is thought to be the neuronal correlate of working memory. For example, a substantial fraction of neurons in the prefrontal cortex elevate their activity persistently during the delay period of spatial working memory tasks in which a monkey has to remember the direction of a visual cue for several seconds. This delay activity is selective to the direction to be memorized.

Problem with the continuous attractor framework.
A common model for spatial working memory is a recurrent network with a ring connectivity, in which the connection strength between two neurons is fine-tuned to depend only on their distance on the ring, and is therefore rotationally invariant. With sufficiently strong and spatially modulated recurrent excitation the network dynamics possess an infinite and continuous set (ring) of attractors. In an attractor the activity is persistent and its profile has the shape of a "bump" localized at an arbitrary location. A transient external stimulus tuned to a specific location, corresponding to the direction to be memorized, selects one of these attractors. After the stimulus is withdrawn, the network remains in this attractor. If the fine-tuning of the interactions is perturbed however, even slightly, the invariance to rotation is broken, and the number of attractors becomes finite and small. Subsequently, after stimulus removal, the persistent bump drifts to the nearest attractor. If this drift is too fast, the network is inappropriate for working memory.

Question: What happens in the presence of synaptic facilitation?

Remembering a 1-dimensional variable

θ is the preferred stimulus of each cell

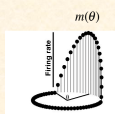


Working memory with fine-tuned synapses:

$$\tau \dot{m}_i = -m_i + \left[I_i + \sum_{j=1}^{N} J_{ij} m_j \right]_+$$

The ring model:

$$J_{ij} = \frac{1}{N}\left(J_0 + 2J_1 \cos(\theta_i - \theta_j)\right)$$



N is the number of minicolumns (in the hundreds)

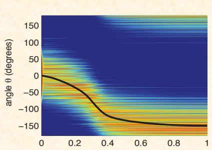
In the presence of a homogeneous input $I_i = I_0$ this has a 1-dimensional family of "bump" attractors centered at the angle ψ

$$m^0(\theta - \psi) = J_1 m_1^0 \left[\cos(\theta - \psi) - \cos(\theta_c)\right]_+$$

Continuous attractor splits into a few discrete attractors in the presence of synaptic heterogeneity.

$$J_{ij} = \frac{1}{N}\left(J_0 + 2J_1 \cos(\theta_i - \theta_j)\right) + \frac{\varepsilon}{\sqrt{N}} n_{ij}$$

Drift of $m(\theta)$ during the delay period is too fast.

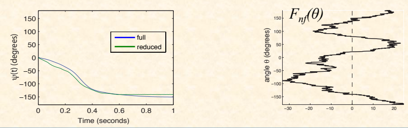


here the synaptic time constant is τ=10ms

Drift dynamics can be understood via a 1-dimensional reduction: $m(\theta,t) = m^0(\theta - \psi(t)) + O(\varepsilon)$

ψ(t) is the center of the "drifting bump."

$$\dot{\psi} = \frac{\varepsilon}{\tau} F_{nf}(\psi) + O(\varepsilon^2), \qquad F_{nf}(\psi) = \frac{1}{\pi m_1^0} \int_{-\theta_c}^{\theta_c} d\theta \sin\theta \int_{-\theta_c}^{\theta_c} d\theta' m^0(\theta') n(\theta + \psi, \theta' + \psi)$$

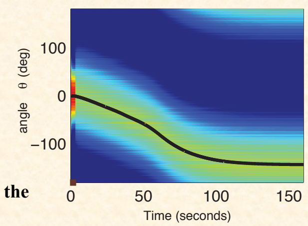


The drift may be significantly slowed in presence of short term plasticity.

$$\begin{cases} \tau \dot{m} = -m + \left[I(\theta) + \int_{-\pi}^{\pi} J(\theta,\theta') u(\theta') x(\theta') m(\theta') d\theta' \right]_+ \\ t_f \dot{u} = -(u - U) + t_f U m (1 - u) \\ t_d \dot{x} = -(x - 1) - t_d m u x \end{cases}$$

t_d=0.1*s*, t_f=1*s*, *U*=0.05, ε=1/2.

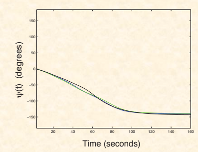
In the regime when facilitation dominates depression, the drift can be slowed down to the scale of many seconds



What controls the speed of the drift?

Simplified model for the facilitating regime U<<1

$$\begin{cases} \tau \dot{m} = -m + \left[I(\theta) + \int_{-\pi}^{\pi} J(\theta, \theta') u(\theta') m(\theta') d\theta' \right]_+ \\ t_f \dot{u} = -(u - U) + t_f U m \end{cases}$$



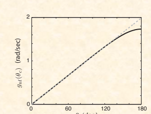
1-dimensional reduction:

$$\dot{\psi} = \varepsilon U F_{fac}(\psi) + O(\varepsilon^2) \qquad F_{fac}(\psi) = \frac{1}{\alpha q_1^0} \int_{-\theta_c}^{\theta_c} d\theta \sin\theta \left(t_f^{-1} + 2m^0(\theta)\right) \int_{-\theta_c}^{\theta_c} d\theta' m^0(\theta') u^0(\theta') n(\theta + \psi, \theta' + \psi)$$

Mean square velocity of the drift:

Computation of mean square velocity without STP:

$$\sqrt{\left\langle \dot{\psi}_{nf}^2 \right\rangle} = \frac{\varepsilon}{\tau\sqrt{N}} g_{nf}(\theta_c) \qquad g_{nf}(\theta_c) = \sqrt{\frac{\theta_c(1 + 2\cos^2\theta_c) - 3\sin\theta_c \cos\theta_c}{\theta_c - \sin\theta_c \cos\theta_c}}$$



Without STP the mean square velocity depends only on the width θ_c of the bump attractor.

Computation of mean square velocity with facilitating synapses:

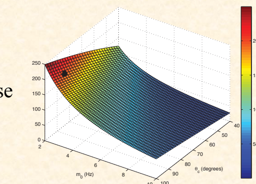
$$\sqrt{< \dot{\psi}_{fac}^2 >} = \frac{\varepsilon U}{\sqrt{N}} g_{fac}(\theta_c, m_0)$$

The drift velocity with facilitating synapses differs by overall factor of τU and also depends on both the magnitude and the width of the "bump"

The drift may be slowed-down by a factor of several hundreds.

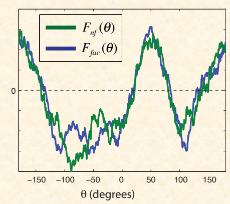
$$r(\theta_c, m_0) = \sqrt{\frac{< \dot{\psi}_{nf}^2 >}{< \dot{\psi}_{fac}^2 >}} = \frac{1}{\tau U} \frac{g_{nf}(\theta_c)}{g_{fac}(\theta_c, m_0)}$$

Prediction: The drift velocity may significantly increase with the increase in average firing rates.

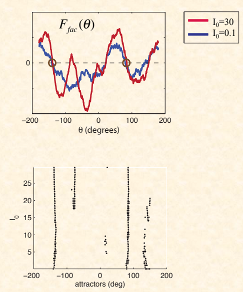


U=0.05, τ=10ms

Attractors in the presence of short-term facilitation closely approximate those without facilitation.



"Strong" attractors are invariant with respect to scaling/ramping of homogenous feedforward input



Conclusions:

- The continuous attractor model is unfeasible for working memory in the absence of fine-tuned synapses and without any additional synaptic mechanism.
- In the presence of the synaptic facilitation the "drift" may be significantly slowed down to the scale of many seconds. This may allow the circuit to implement a working memory.
- Average firing rates may control the velocity of the "drift" in the presence of synaptic facilitation.
- Persistent attractors are a feature of the synaptic heterogeneity n_{ij} and are robust to the fluctuations in the homogeneous feedforward input and the particular details of STP.