Structured memory spaces arising from recurrent network activity

Vladimir Itskov, Columbia University
Two complementary pictures for what determines the activity of individual neurons

\[ x_i \approx F_i(s) \]

\[ \dot{x} = -x + [Jx + b]_+ \]

Input: \( b=(b_1,..b_N) \)

RECEPTIVE FIELDS

RECURRENT NETWORK
The set of possible activity patterns

Both pictures suggest constraints on allowed activity patterns...
PLAN OF THE TALK

The set of possible activity patterns

\[ p_1 \quad p_2 \quad p_3 \]

Today

Stimulus space

Hahnloser & Seung (2002-3)


Recurrent network
A toy model for receptive field picture, inspired by hippocampal place-fields

\[ x_i = F_i(s) \]

**Cell groups** = groups of neurons that are persistently co-active due to a stimulus that lies in the intersection of their corresponding receptive fields.
WHAT IF THE PURPOSE OF THE RECURRENT NETWORK IS TO IMPLEMENT RECEPTIVE FIELD CONSTRAINTS?

Receptive field picture:

Cells 1, 2, 3 can persistently fire together
Cells 4, 5, 6 can persistently fire together
Cells 3 and 6 are forbidden to fire together
WHAT IF THE PURPOSE OF THE RECURRENT NETWORK IS TO IMPLEMENT RECEPTIVE FIELD CONSTRAINTS?

Two ways a network can forbid activity patterns that “do not make sense”

1. Feed-forward inputs $b$ already reflect the same exact stimulus space, and are already appropriately constrained.

Cells 1, 2, 3 can persistently fire together
Cells 4, 5, 6 can persistently fire together
Cells 3 and 6 are forbidden to fire together

\[ \dot{x} = -x + [Jx + b]_+ \]
WHAT IF THE PURPOSE OF THE RECURRENT NETWORK IS TO IMPLEMENT RECEPTIVE FIELD CONSTRAINTS?

Receptive field picture:

Cells 1, 2, 3 can persistently fire together
Cells 4, 5, 6 can persistently fire together
Cells 3 and 6 are forbidden to fire together

Two ways a network can forbid activity patterns that “do not make sense”

1. Feed-forward inputs $b$ already reflect the same exact stimulus space, and are already appropriately constrained.

2. Recurrent network forbids nonsensical activity patterns, so the network can tolerate a wide range of inputs and still reflect receptive field constraints.

\[ \dot{x} = -x + [Jx + b]_+ \]
Recurrent network


The set of possible activity patterns

$\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3$

Stimulus space

Today

Hahnloser & Seung (2002-3)

Recurrent network
The set of possible activity patterns

$\{p_1, p_2, p_3\}$

Hahnloser & Seung (2002-3)

Recurrent network
A toy model for a recurrent network

\[ \dot{x} = -x + \left[ Jx + b \right]_+ \]

Input: \( b=(b_1, \ldots, b_N) \)

N.B. The set of all b’s is not the stimulus space! They are the inputs to each neuron in the recurrent network, and range over all of \( \mathbb{R}^N \) whereas the stimulus space is generally low-dimensional.
Forbidden sets

\[ \dot{x} = -x + [Jx + b]_+ \]

A group of cells is called forbidden if there does not exist any constant input \( b \) such that there exists a steady state in which these and only these cells fire.

**Key observation**: forbidden sets exist!

A set of cells $c=(c_1,..,c_k)$ is called *permitted* if it is not forbidden.

I.e., there exists a constant input vector $b$ and an initial condition such that the system converges to the stable steady state whose only activated cells are exactly $c$.

$\dot{x} = -x + [Jx + b]_+$

If there are no strong recurrent connections, everything is permitted.

Input: $b=(b_1, \ldots, b_N)$

$$\dot{x} = -x + \left[ Jx + b \right]_+$$
Some useful observations:

\[ \dot{x} = -x + \left[ Jx + b \right]_+ \]

If the synaptic matrix \( J \) is symmetric, then:

- A set of cells \( c \) is permitted if and only if the appropriate submatrix \( J_c \) has \( \max \lambda_i(J_c) < 1 \)

- A cell group \( c \) is forbidden if the corresponding sub-network is \textit{unstable}, i.e. the appropriate submatrix satisfies \( \max \lambda_i(J_c) \geq 1 \)

\textit{We can determine the set of permitted sets directly from the matrix \( J \).}

- Every subset of a permitted set is permitted; every superset of forbidden set is forbidden. This makes the set of all permitted sets a “simplicial complex” \( K \).

Every subset of a permitted set is permitted; every superset of a forbidden set is forbidden.

This means the set of permitted sets is a “simplicial complex.”
Some generalization for $J$ not symmetric:

**Lemma (V.I.):** A set $c$ of cells is permitted iff the submatrix $J_c$ corresponding to the sub-network $c$ is stable. I.e., if $\max \Re \lambda_i(J_c) < 1$.

A caveat: in this case, the set $\text{perm}(J)$ is not a simplicial complex. However, we can still consider the smallest simplicial complex containing $\text{perm}(J)$. 
The set of possible activity patterns

Set of permitted sets

Hahnloser & Seung (2002-3)

Recurrent network

Synaptic weight matrix J
The set of possible activity patterns

$\mathbf{p}_1$  $\mathbf{p}_2$  $\mathbf{p}_3$

Cell groups

Global topological features

Stimulus space

Homology groups $H_d(S)$ keep track of $(d+1)$-dimensional “holes” (with $d$-dimensional boundaries) in $S$, and thus reflect topological features of the stimulus space.
The set of possible activity patterns

Homology groups of simplicial complex defined by cell groups

Stimulus space

Homology groups of stimulus space

The set of possible activity patterns

$\text{Homology groups of simplicial complex defined by cell groups permitted sets}$


Recurrent network

Hahnloser & Seung (2002-3)
Recurrent network

The set of possible activity patterns

\[ \begin{align*}
  p_1 \\
  p_2 \\
  p_3
\end{align*} \]

Stimulus space

Putting it all together...


Hahnloser & Seung (2002-3)

Recurrent network
The set of possible activity patterns

$p_1$ $p_2$ $p_3$

Question: can any set of permitted sets be realized via this recurrent network model?
Some technical definitions:

Maximal permitted sets of K:

\[ K_{n}^{\text{max}} \overset{\text{def}}{=} \{ c \in K_n \mid \nexists \tilde{c} \in \bigcup_{n' = n+1}^{\infty} K_{n'}, \text{ such that } \tilde{c} \supset c \} \]

Minimal forbidden sets of K:

\[ \bar{K}_{n}^{\text{min}} \overset{\text{def}}{=} \{ c \in \bar{K}_n \mid \nexists \tilde{c} \in \bigcup_{n' = 0}^{n-1} \bar{K}_{n'}, \text{ such that } \tilde{c} \subset c \} \]
Given a set of permitted sets $K$, define an energy function on symmetric matrices $J$ as:

$$E_K(J) = \sum_{c \in K_{\text{max}}} \left[ -1 + \lambda_{\text{max}}(J^c) \right]_+ + \sum_{\bar{c} \in \bar{K}_{\text{min}}} \left[ 1 - \lambda_{\text{max}}(J^{\bar{c}}) \right]_+$$

Note that $E_K(J) \geq 0$, and $E_K(J)=0$ if and only if $\text{perm}(J)=K$.

Therefore learning $K$ involves finding $J$ such that $E_K(J)=0$, i.e. minimizing $E_K(J)$. 
We can use the gradient of the energy function as the learning rule:

\[ \dot{J} = -\nabla E_K(J) \]

The gradient of \( E_K(J) \) can be written in terms of “highest eigenvectors” \( v_c \) corresponding to appropriate cell groups \( c \).

\[ \nabla E_K(J) = \sum_{c \in K^{\text{max}}} \theta(-1 + \lambda_{\text{max}}(J^c))(v_c v_c^T) - \sum_{\bar{c} \in K^{\text{min}}} \theta(1 - \lambda_{\text{max}}(J^{\bar{c}}))(v_{\bar{c}} v_{\bar{c}}^T) \]

where \( \theta \) is the Heaviside step-function, and \( v_c v_c^T \) is the orthonormal projector to the “maximal eigenspace.”
Learning rule

We can use the gradient of the energy function as the learning rule:

\[ \dot{J} = -\nabla E_K(J) \]

The gradient of \( E_K(J) \) can be written in terms of “highest eigenvectors” \( v_c \) corresponding to appropriate cell groups \( c \).

\[ \nabla E_K(J) = \sum_{c \in K^{\text{max}}} \theta(-1 + \lambda_{\text{max}}(J^c))(v_cv_c^T) - \sum_{\bar{c} \in K^{\text{min}}} \theta(1 - \lambda_{\text{max}}(J^{\bar{c}}))(v_{\bar{c}}v_{\bar{c}}^T) \]

where \( \theta \) is the Heaviside step-function, and \( v_c v_c^T \) is the orthonormal projector to the “maximal eigenspace.”

It works!
The set of possible activity patterns

Recurrent network

Stimulus space
What is all of this good for, anyway?

We can find conditions on the recurrent network $J$ such that the set of all permitted sets satisfies (and thus implements) the stimulus space constraints.
An Example – The Ring Model

\[ \dot{x}(\theta) = -x + \left[ \int_{-\pi}^{\pi} J(\theta - \theta') x(\theta') d\theta' + b(\theta) \right]_+ \]

\[ J_{ij} = \frac{1}{N} \left( j_0 + j_1 \cos(\theta_i - \theta_j) \right) \]

An Example – The Ring Model

Here we’d like to see under what conditions the structure of the space of permitted sets is the same as the structure of the space of orientations (a circle). It does not generally need to be true.

\[ J_{ij} = \frac{1}{N} \left( j_0 + j_1 \cos(\theta_i - \theta_j) \right) \]
An Example – The Ring Model

Here we’d like to see if the structure of the space of permitted sets has the same as the structure of the space of orientations (a circle). It does not generally need to be the same.

\[ J_{ij} = \frac{1}{N} \left( j_0 + j_1 \cos(\theta_i - \theta_j) \right) \]
How it is actually done…

For each of these sets of permitted sets $K$, we can write the energy as a function of the parameters of $J$.

Recall: $E_K(J)=0$ if and only if $\text{perm}(J)=K$. 

$E_K(J) = \sum_{c \in K^{\text{max}}} [-1 + \lambda_{\text{max}}(J^c)]_+ + \sum_{\overline{c} \in \overline{K}^{\text{min}}} [1 - \lambda_{\text{max}}(J^{\overline{c}})]_+$
A phase diagram for the ring model

\[ \dot{x} = -x + \left[ Jx + b \right]_+ \]

\[ J_{ij} = \frac{1}{N} \left( j_0 + j_1 \cos(\theta_i - \theta_j) \right) \]
**Question:** What if the network (i.e. J) is random?
**Question:** What if the network (i.e. J) is random?

**Answer:** Ugly, high-dimensional topology.

**The lesson:** If permitted sets reflect a “nice” stimulus space (like in hippocampus), we should get some strong constraints on network connectivity.
PLAN OF THE TALK

The set of possible activity patterns

$p_1$  $p_2$  $p_3$

Stimulus space

Recurrent network


Today

Hahnloser & Seung (2002-3)
Acknowledgements

Supported by the Swartz Foundation.
This is joint work with Carina Curto.

Columbia Center for Theoretical Neuroscience