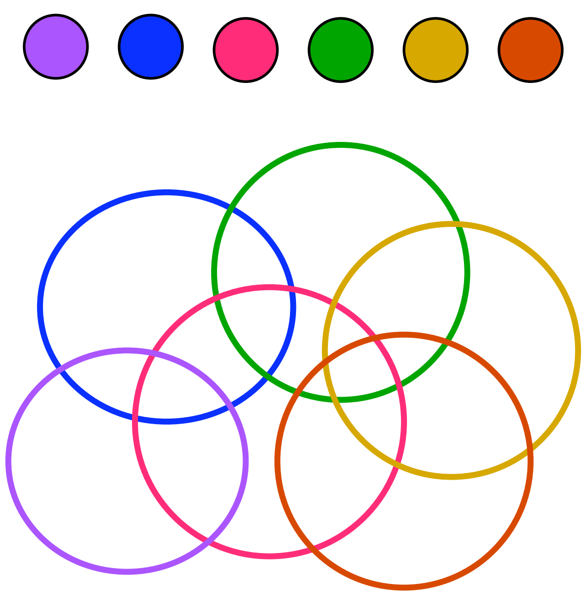
Structured memory spaces arising from recurrent network activity

Vladimir Itskov, Columbia University

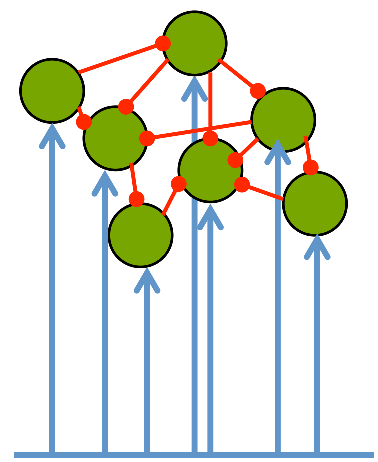
Two complementary pictures for what determines the activity of individual neurons

$x_i \approx F_i(s)$



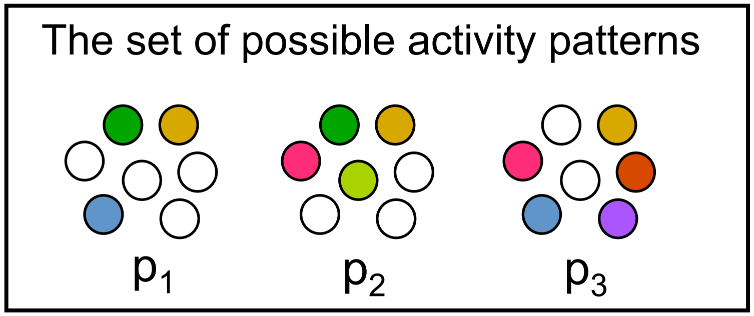
RECEPTIVE FIELDS

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



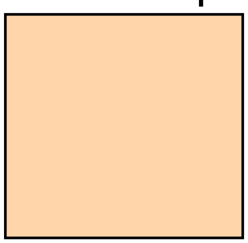
Input: $b=(b_1,..b_N)$

RECURRENT NETWORK

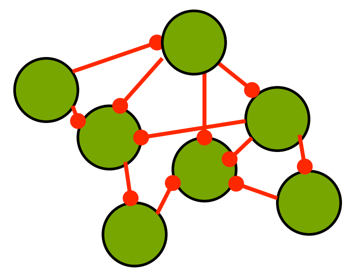


Both pictures suggest constraints on allowed activity patterns…

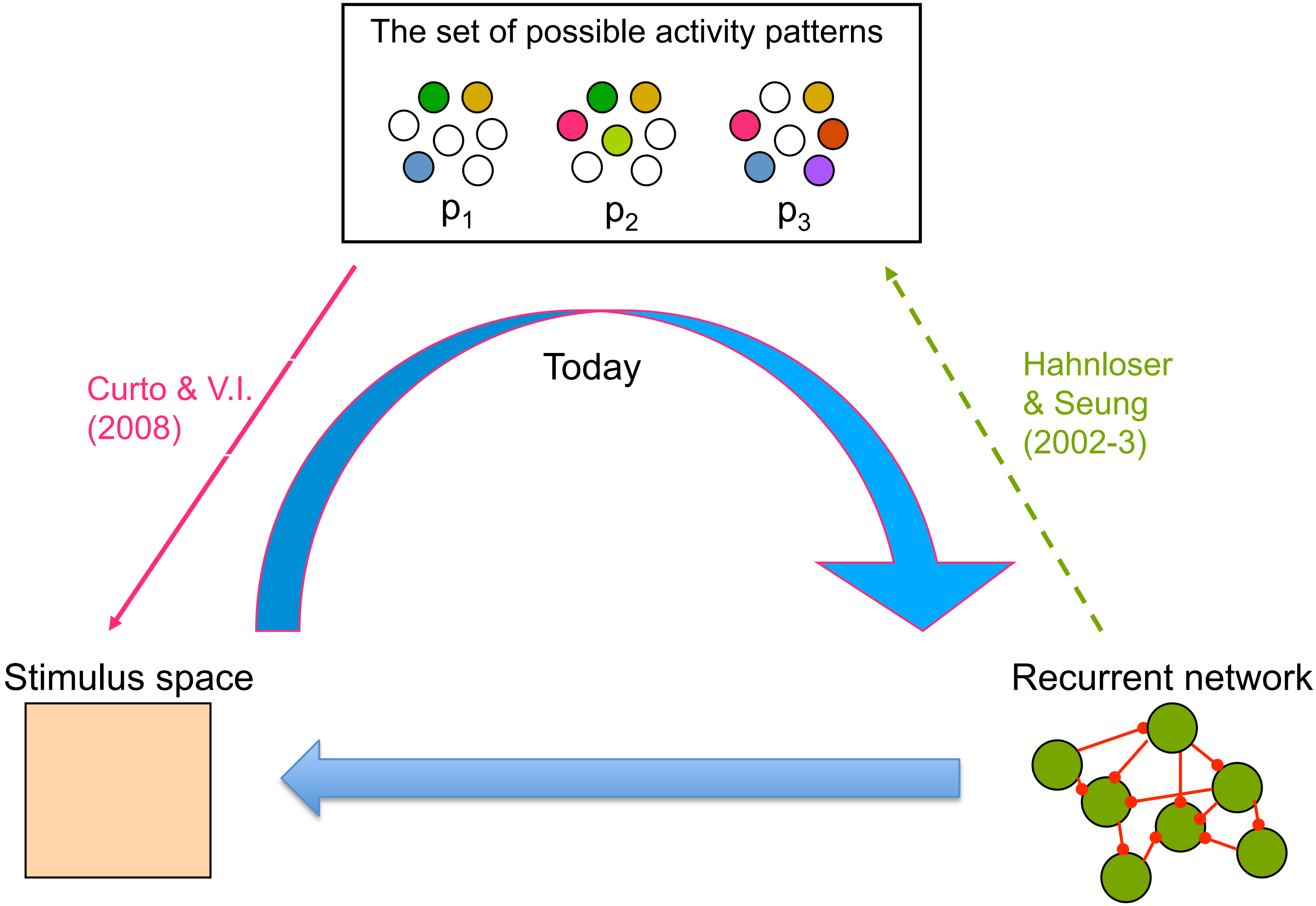
Stimulus space



Recurrent network

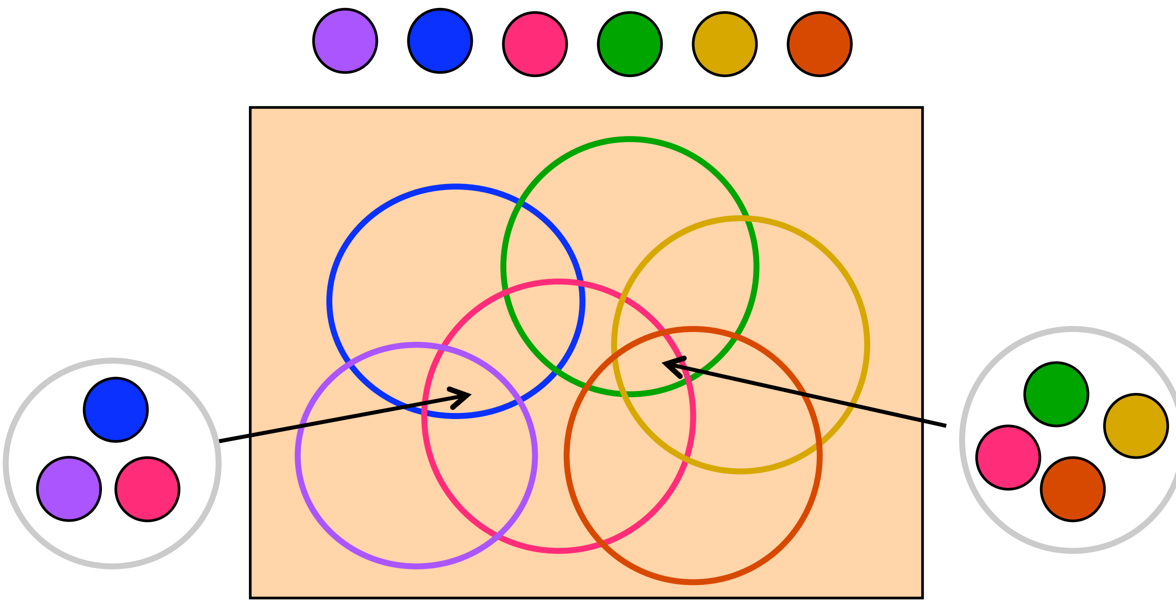


PLAN OF THE TALK



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

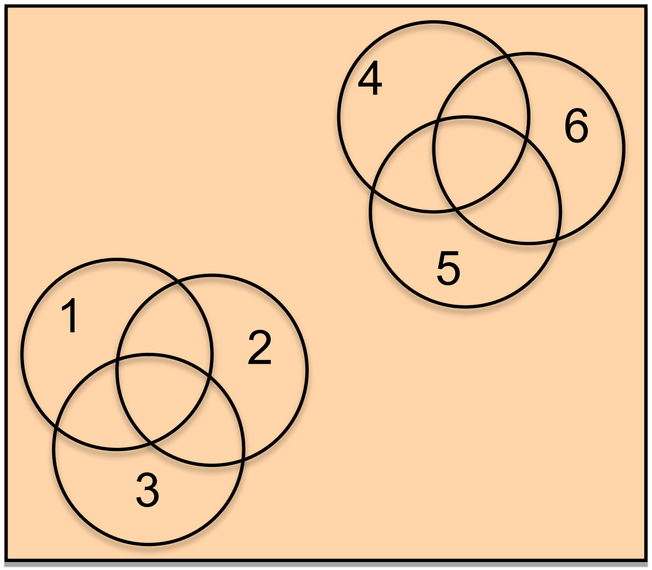
$x_i = F_i(s)$



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:

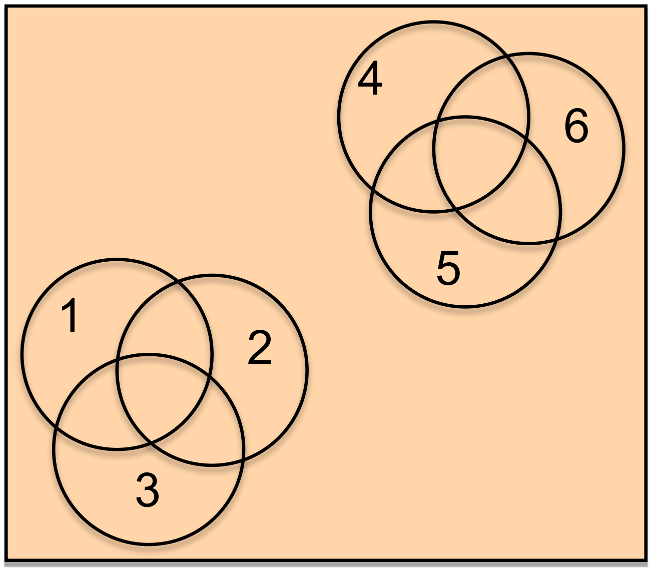


Cells 1,2,3 can persistently fire together
Cells 4,5,6 can persistently fire together

Cells 3 and 6 are *forbidden* to fire together

WHAT IF THE PURPOSE OF THE RECURRENT NETWORK IS TO IMPLEMENT RECEPTIVE FIELD CONSTRAINTS?

Receptive field picture:

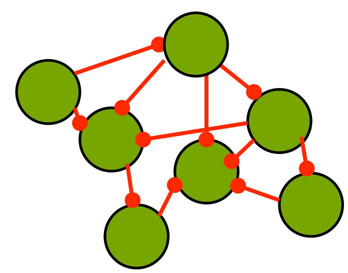


Two ways a network can forbid activity patterns that "do not make sense"

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

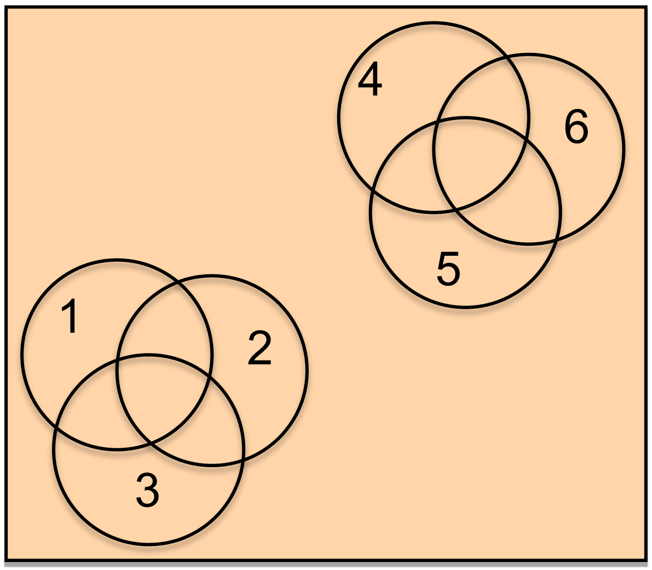
Cells 3 and 6 are *forbidden* to fire together



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

WHAT IF THE PURPOSE OF THE RECURRENT NETWORK IS TO IMPLEMENT RECEPTIVE FIELD CONSTRAINTS?

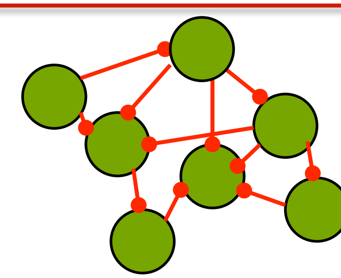
Receptive field picture:



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.

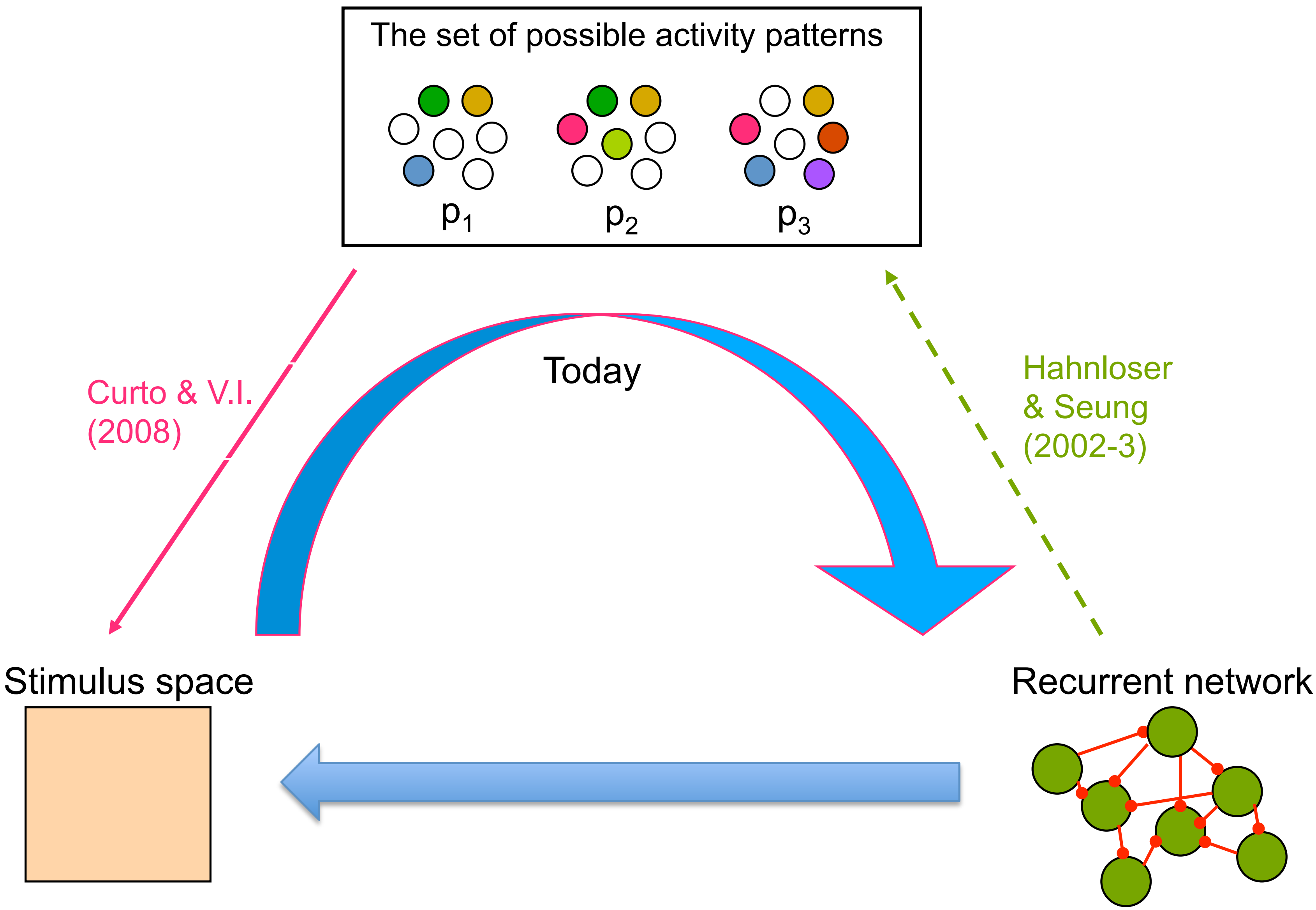
2. Recurrent network forbids nonsensical activity patterns, so the network can tolerate a wide range of inputs and still reflect receptive field constraints.



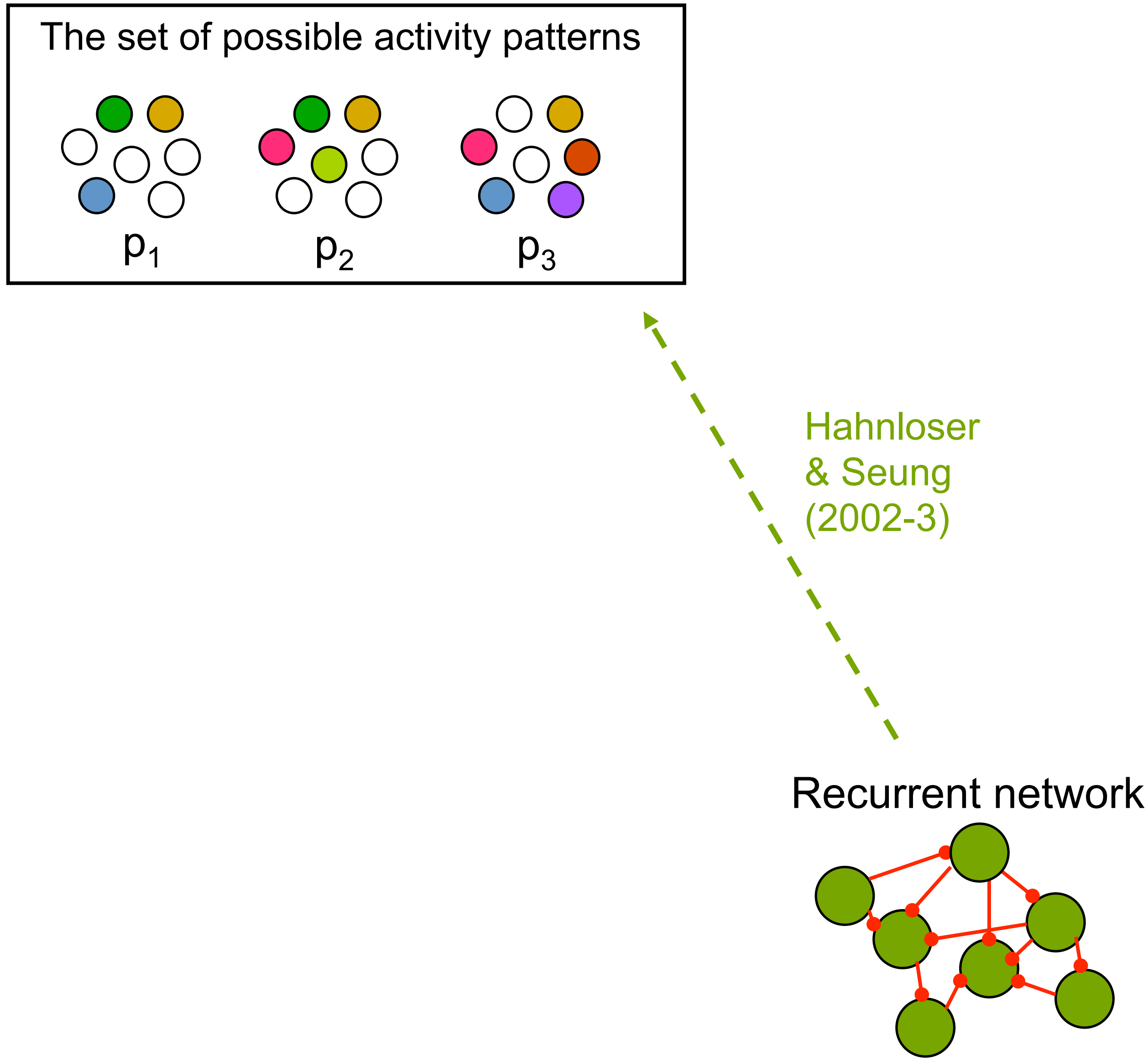
$$\dot{x} = -x + [Jx + b]_+$$

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



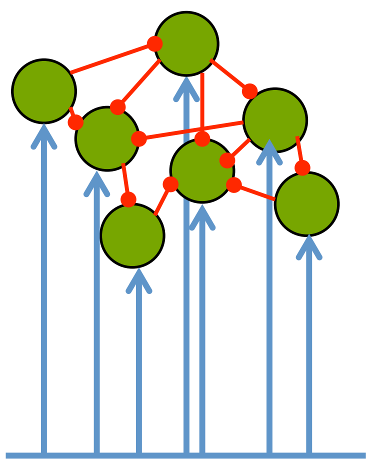
The set of possible activity patterns
p1
p2
p3
Curto & V.I.
(2008)
Today
Hahnloser
& Seung
(2002-3)
Stimulus space
Recurrent network



The set of possible activity patterns
p1
p2
p3
Hahnloser
& Seung
(2002-3)
Recurrent network

A toy model for a recurrent network

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

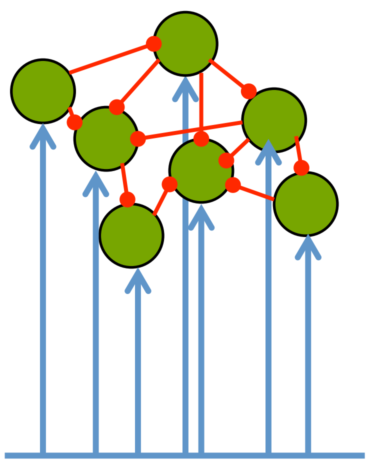


Input: $b=(b_1,..b_N)$

N.B. <u>The set of all b's is not the stimulus space</u>! They are the inputs to each neuron in the recurrent network, and range over all of R^N whereas the stimulus space is generally low-dimensional.

Forbidden sets

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



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

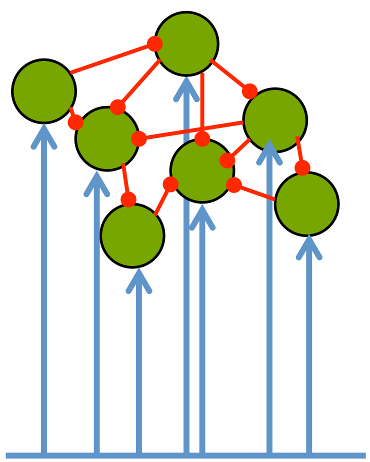
A group of cells is called *forbidden* if there *does not exist any* constant input b such that there exists a steady state in which these and only these cells fire.

<u>Key observation</u>: forbidden sets exist!

Hahnloser, Seung, Slotine. Neural Comput. (2003).

Permitted sets

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



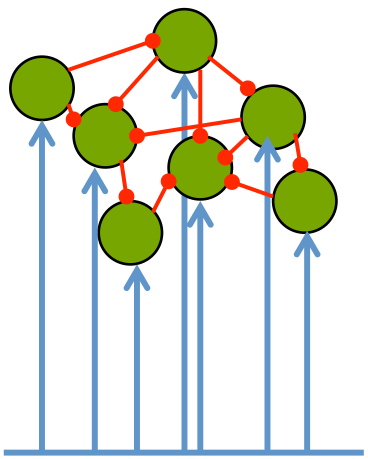
Input: $b=(b_1,..b_N)$

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.

Hahnloser, Seung, Slotine. Neural Comput. (2003).

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



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

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

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 *unstable,* i.e. the appropriate submatrix satisfies $\max \lambda_i(J_c) \geq 1$

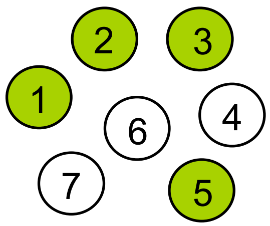
  We can determine the set of permitted sets directly from the matrix J.

- 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.

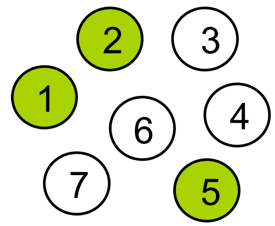
Xie, Hahnloser, and Seung. Neural Comput. (2002), Hahnloser, Seung, Slotine. Neural Comput.(2003).

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

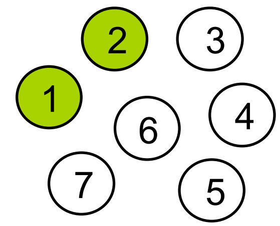
permitted



[1 2 3 5]



[1 2 5]

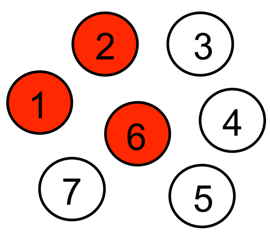


[1 2]

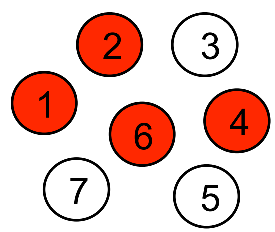
This means the set of permitted sets is a "*simplicial complex.*"

---

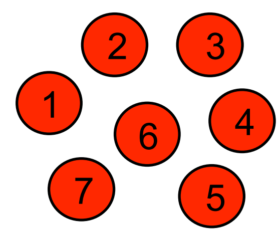
forbidden



[1 2 6]



[1 2 4 6]

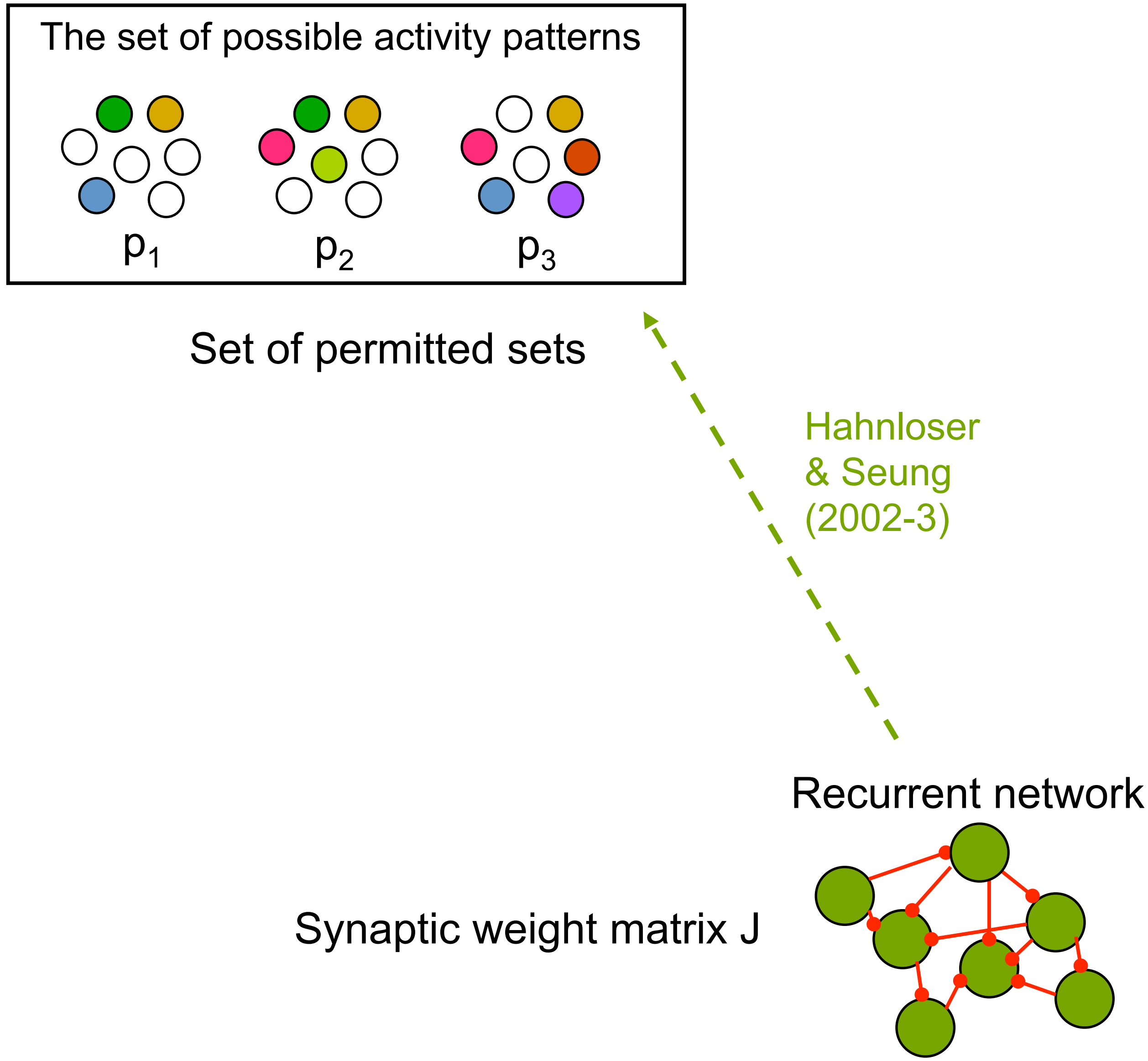


[1 2 3 4 5 6]

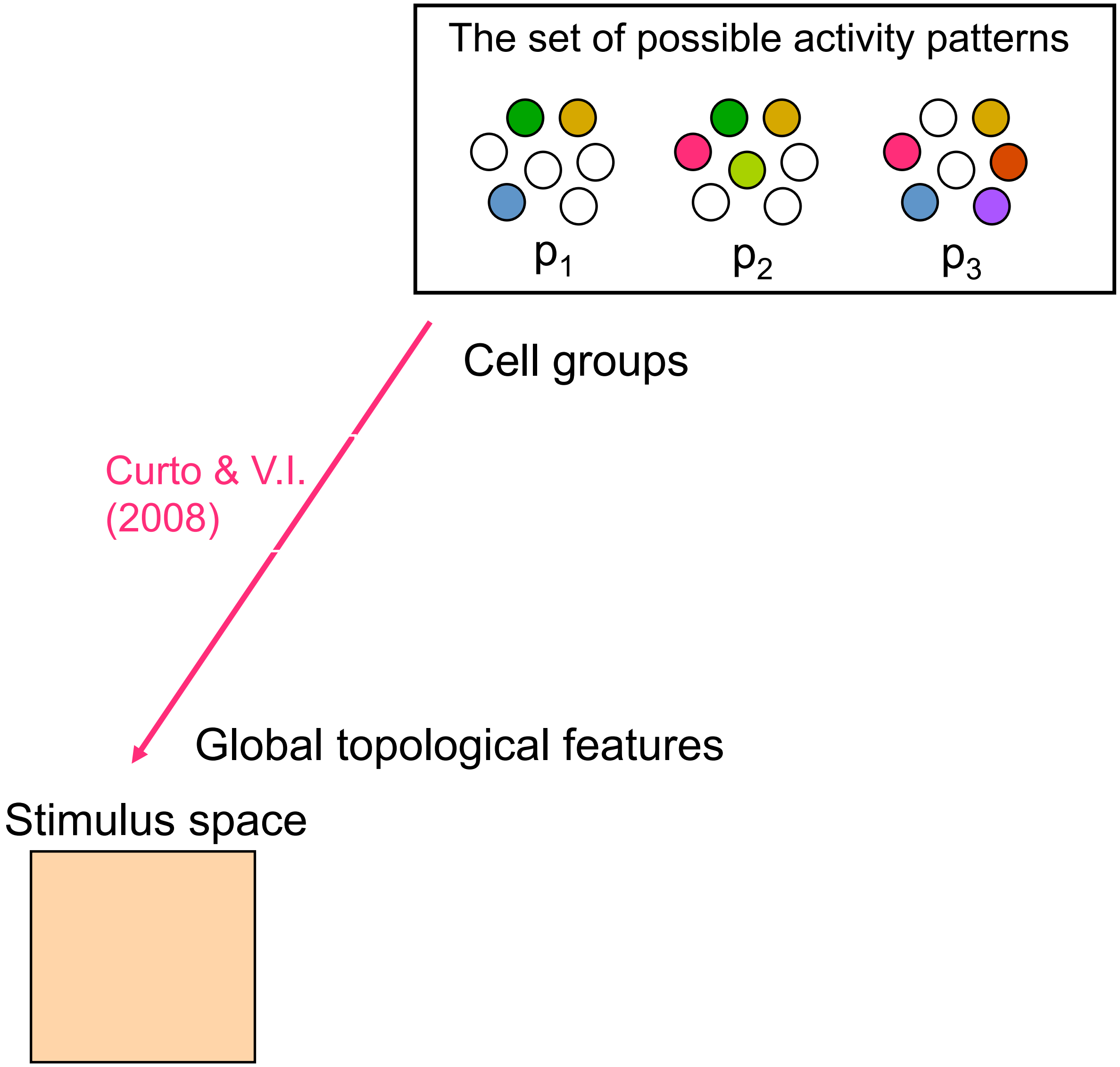
Some generalization for J not symmetric:

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 $\max \operatorname{Re} \lambda_i(J_c) < 1$.

A caveat: in this case, the set perm(J) is not a simplicial complex. However, we can still consider the smallest simplicial complex containing perm(J).



The set of possible activity patterns
p_1
p_2
p_3
Set of permitted sets
Hahnloser
& Seung
(2002-3)
Recurrent network
Synaptic weight matrix J


The set of possible activity patterns
p1
p2
p3
Cell groups
Curto & V.I.
(2008)
Global topological features
Stimulus space

An algorithm: cell groups to topology

The set of all cell groups that are persistently co-activated

↓

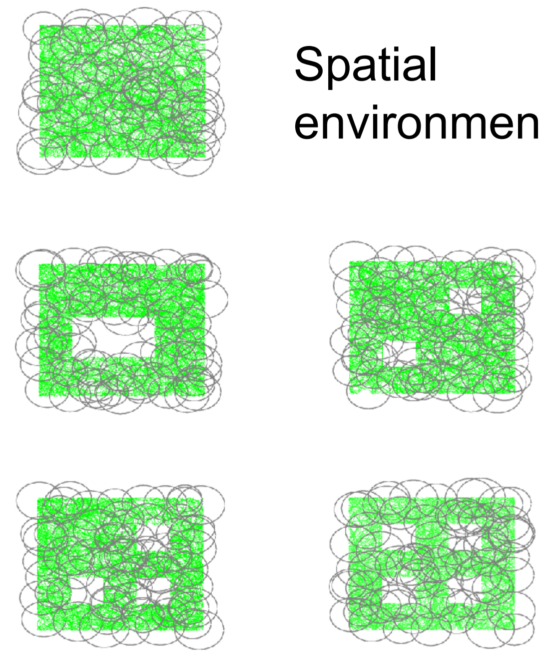
Receptive field intersection information (simplicial complex K)

↓

Simplicial Homology $H_d(K)$ → Homology $H_d(S)$ of the underlying stimulus space S.

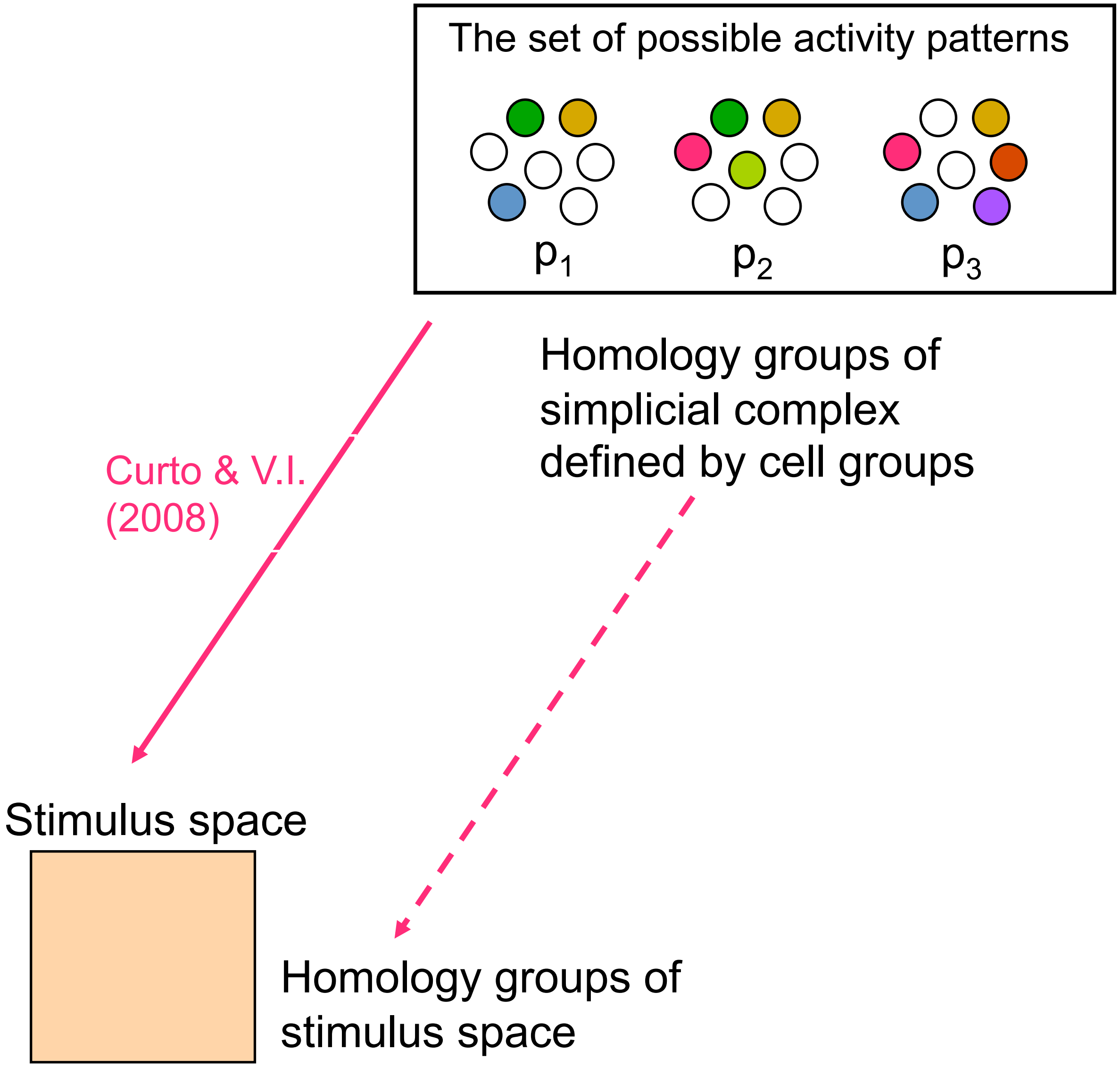
=

Theorem in algebraic topology

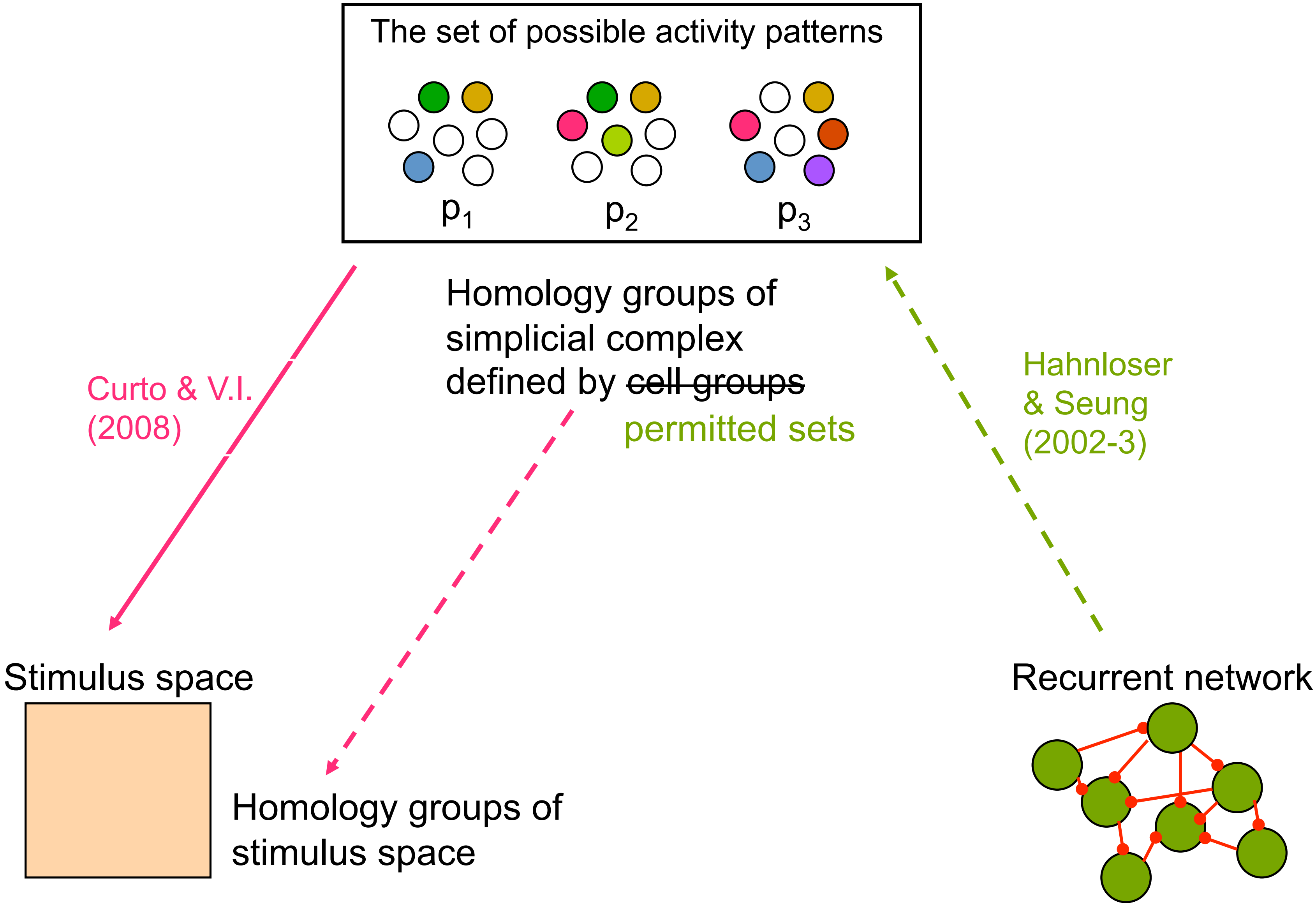


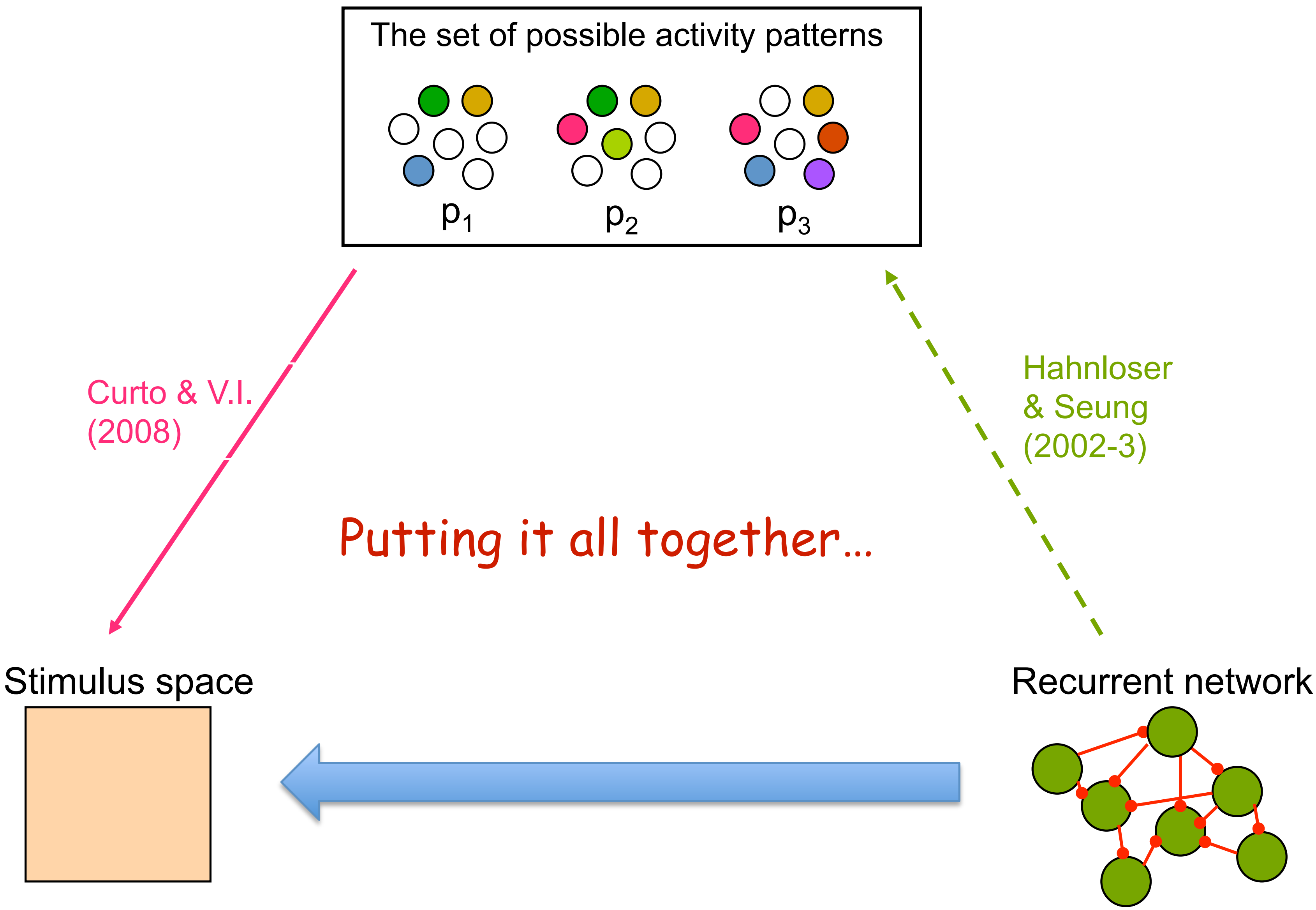
Spatial environments

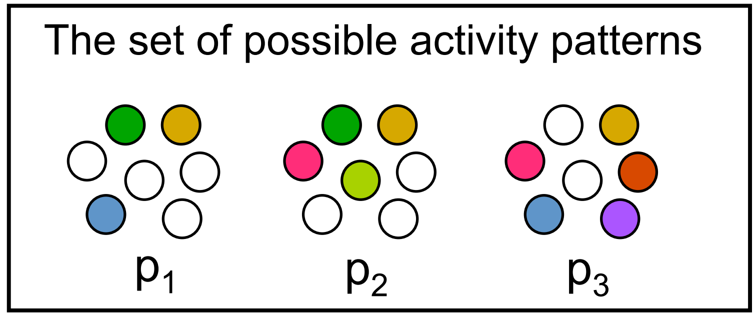
Homology groups $H_d(S)$ keep track of (d+1)-dimensional “holes” (with d-dimensional boundaries) in S, and thus reflect topological features of the stimulus space.

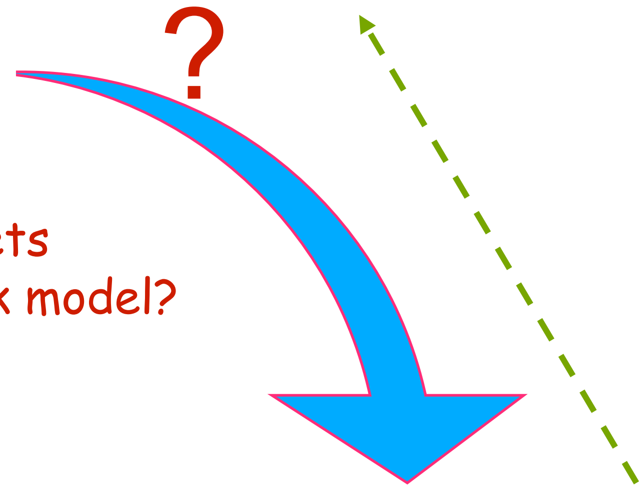


The set of possible activity patterns
p_1
p_2
p_3
Homology groups of
simplicial complex
defined by cell groups
Curto & V.I.
(2008)
Stimulus space
Homology groups of
stimulus space


The set of possible activity patterns
p_1
p_2
p_3
Homology groups of
simplicial complex
defined by ~~cell groups~~
permitted sets
Curto & V.I.
(2008)
Hahnloser
& Seung
(2002-3)
Stimulus space
Homology groups of
stimulus space
Recurrent network

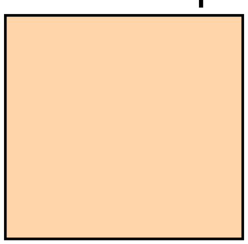

The set of possible activity patterns
p_1
p_2
p_3
Curto & V.I.
(2008)
Hahnloser
& Seung
(2002-3)
Putting it all together...
Stimulus space
Recurrent network



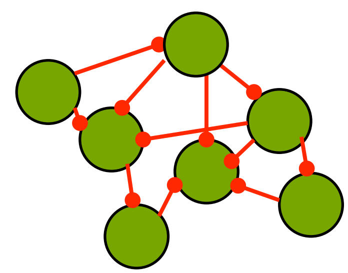


Question: can any set of permitted sets
be realized via this recurrent network model?

Stimulus space

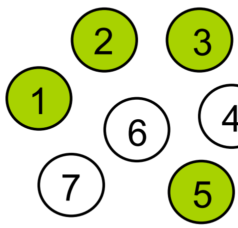


Recurrent network



Some technical definitions:

maximal
permitted set

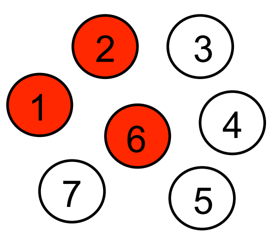


Maximal permitted sets of K:

$$\mathrm{K}_n^{\max} \stackrel{\text{def}}{=} \{c \in K_n \mid \nexists \tilde{c} \in \textstyle\bigcup_{n'=n+1}^{\infty} K_{n'},\ \text{such that } \tilde{c} \supset c\}$$

---

minimal
forbidden set



Minimal forbidden sets of K:

$$\bar{K}_n^{\min} \stackrel{\text{def}}{=} \{c \in \bar{K}_n \mid \nexists \tilde{c} \in \bigcup_{n'=0}^{n-1} \bar{K}_{n'},\ \text{such that } \tilde{c} \subset c\}$$

Energy function

Given a set of permitted sets K, define an energy function on symmetric matrices J as:

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

Note that $E_K(J)$>=0, and $E_K(J)$=0 if and only if perm(J)=K.

Therefore learning K involves finding J such that $E_K(J)$=0 , i.e. minimizing $E_K(J)$.

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.

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

where θ 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:

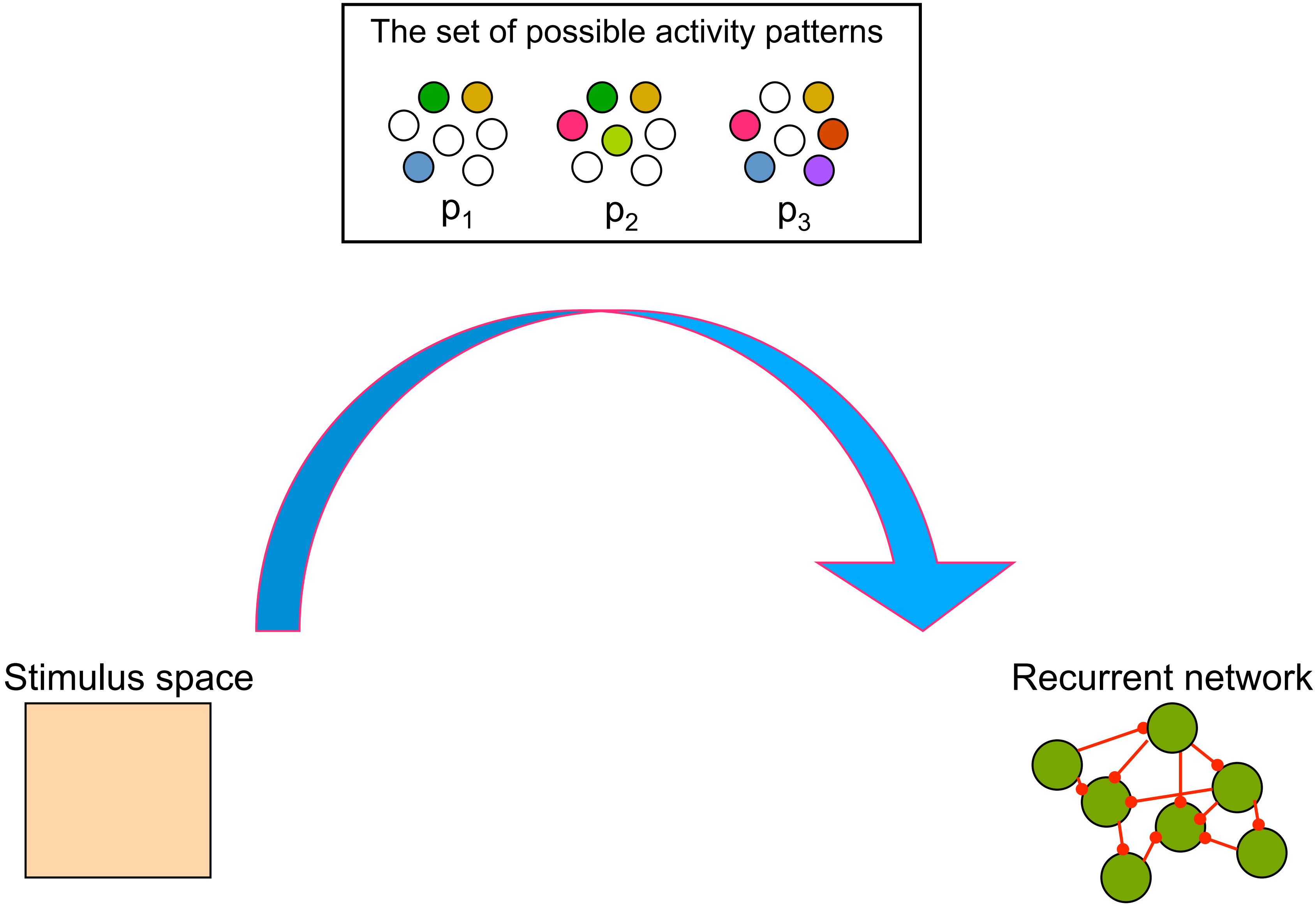
$$\dot{J} = -\nabla E_K(J)$$

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

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

where θ is the Heaviside step-function, and $v_c v_c{}^T$ is the orthonormal projector to the "maximal eigenspace."

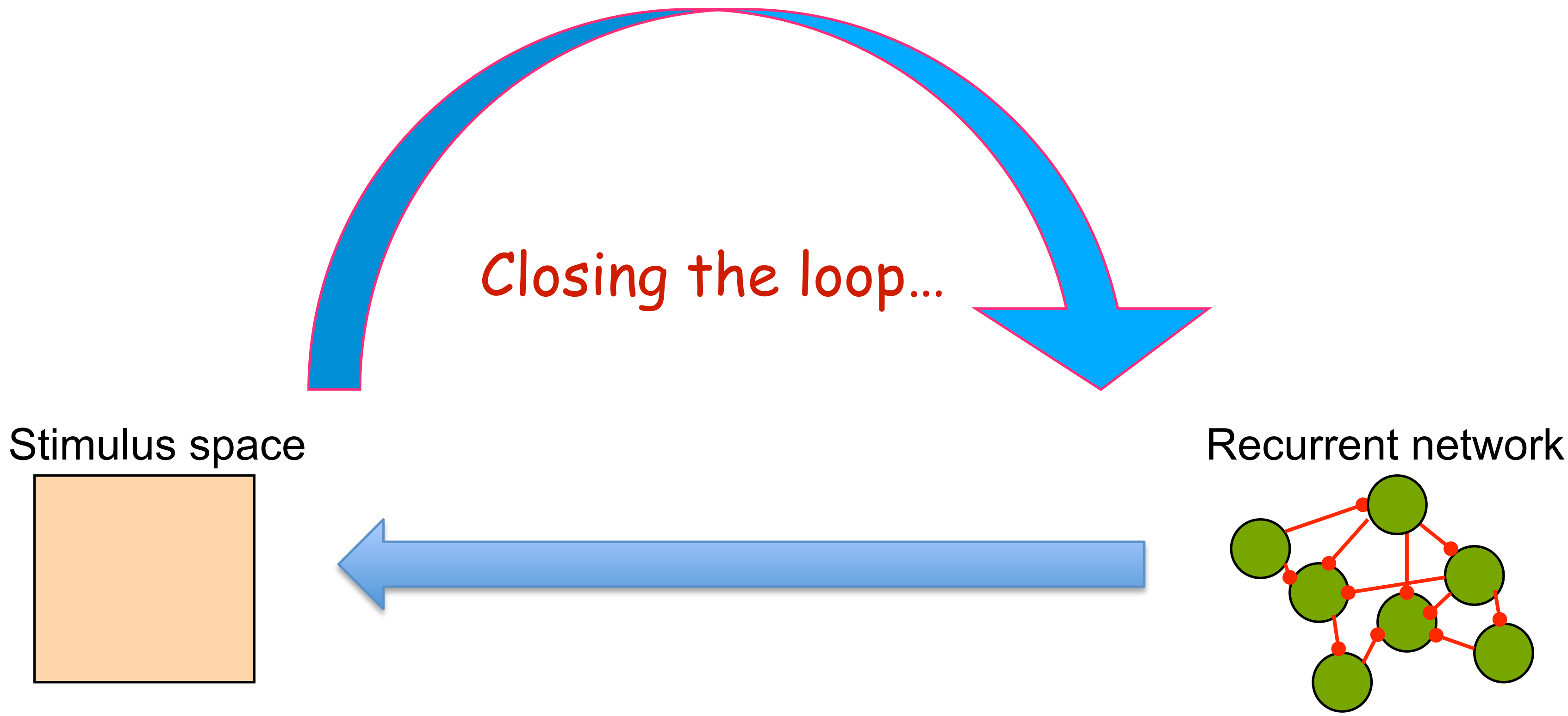
It works!



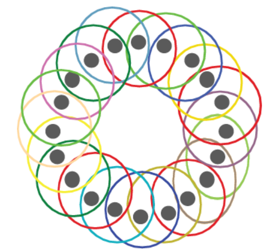
The set of possible activity patterns
p1
p2
p3
Stimulus space
Recurrent network

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



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

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

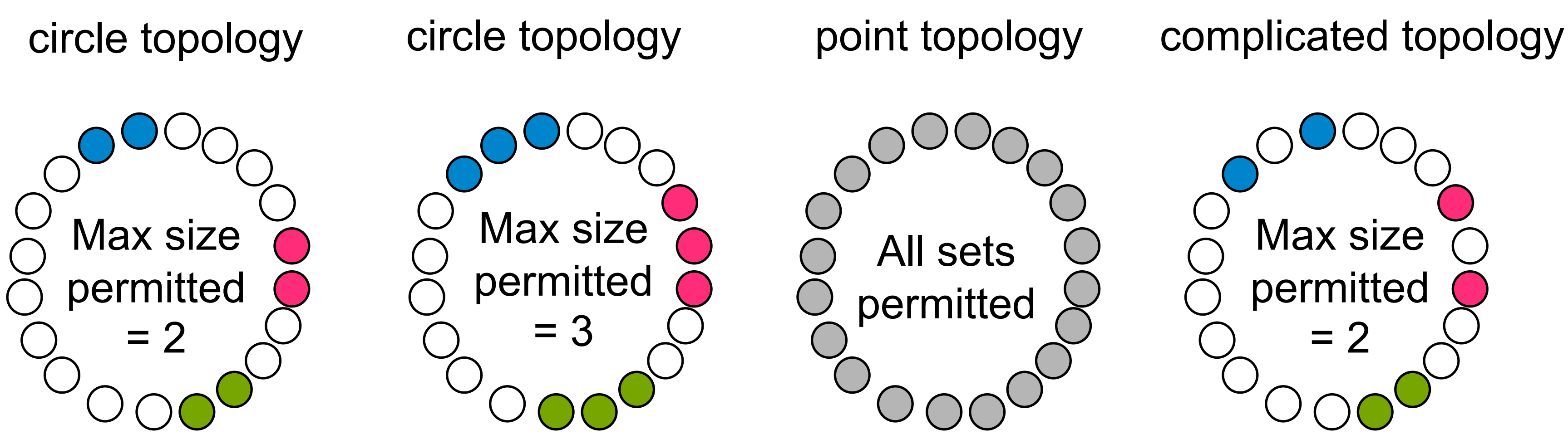
All-to-all connections!

Ben-Yishai, Bar-Or & Sompolinsky, PNAS 1995.

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.**

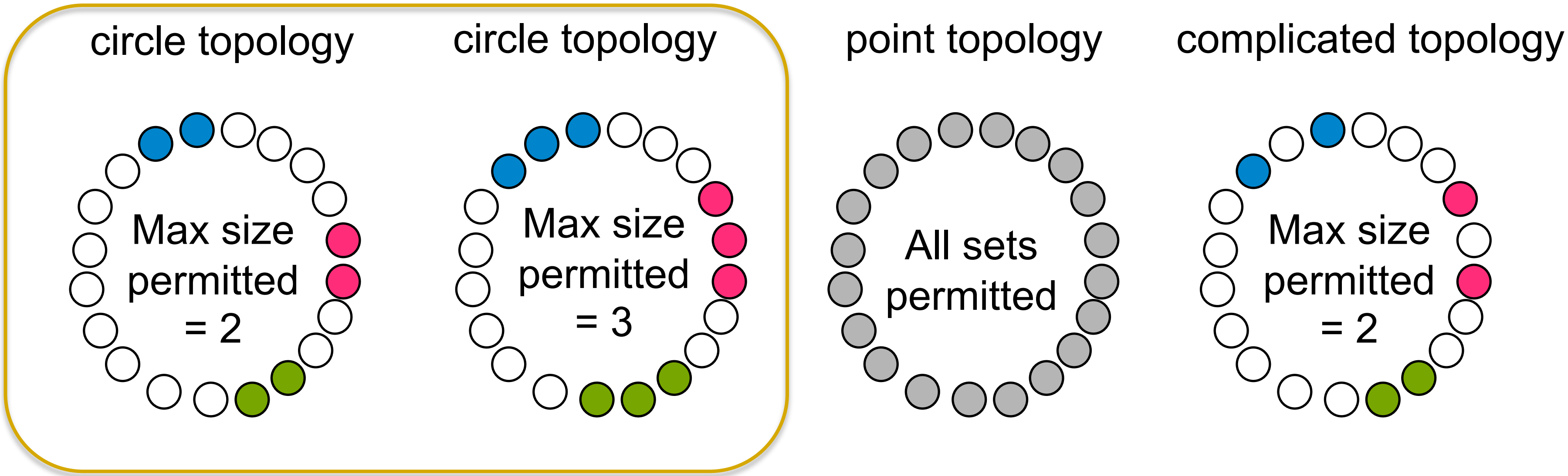
$$J_{ij} = \frac{1}{N}\left(j_0 + j_1 \cos(\theta_i - \theta_j)\right)$$



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.**

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

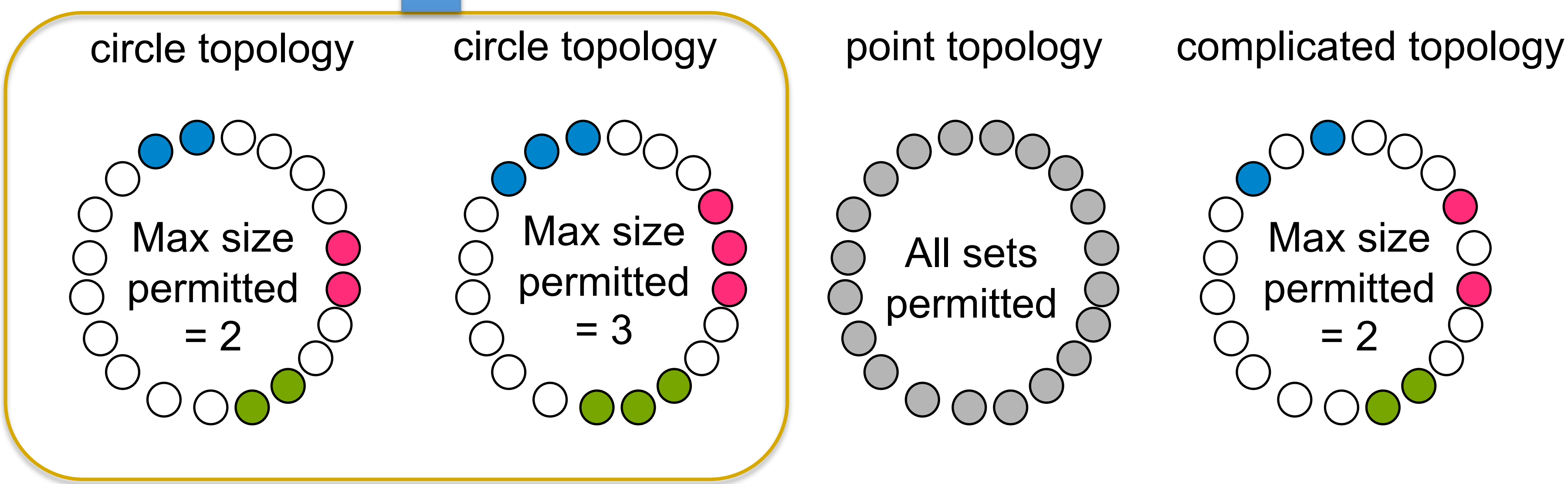


How it is actually done…

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

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

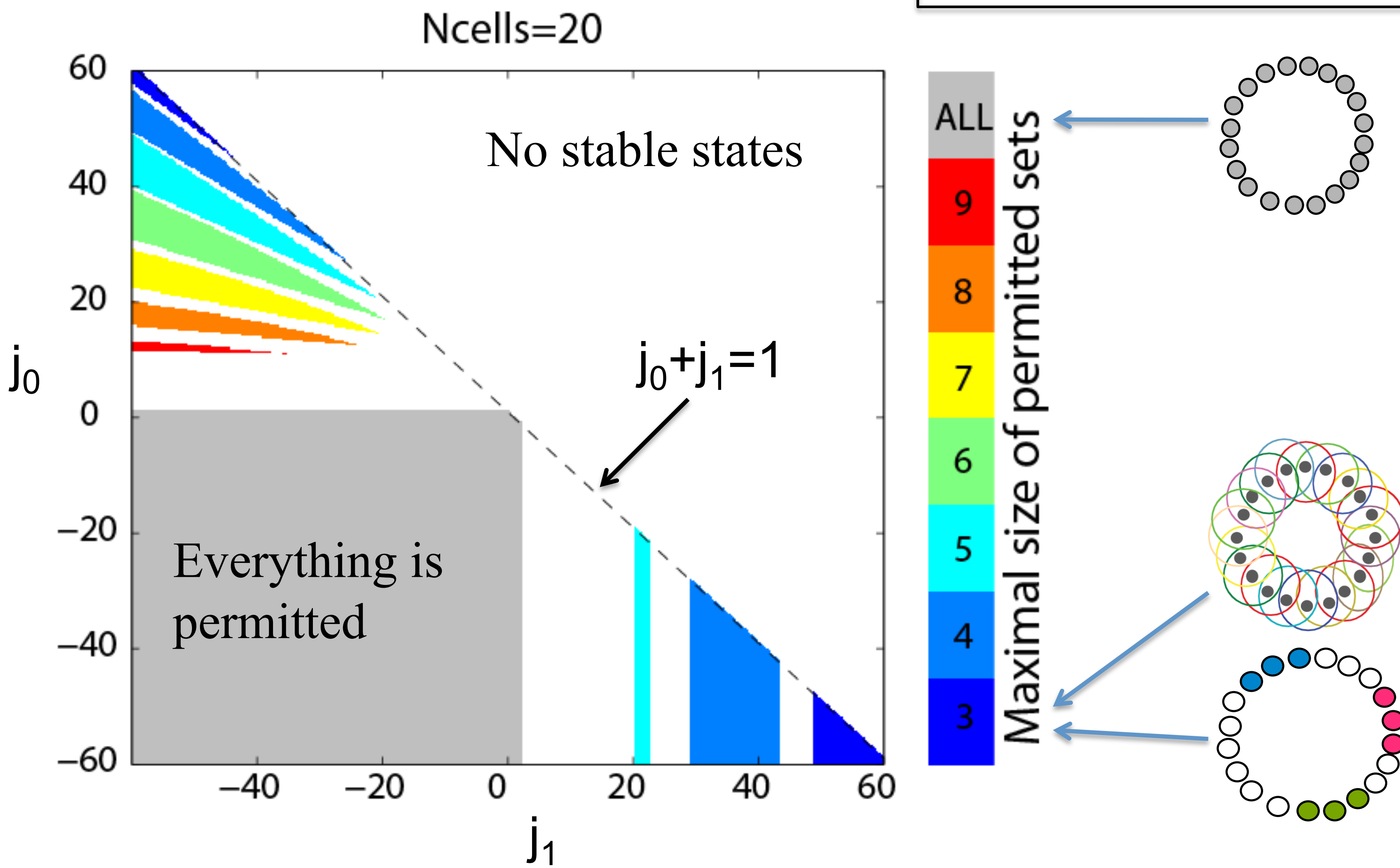
Recall: $E_K(J)=0$ if and only if perm(J)=K.



A phase diagram for the ring model

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

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



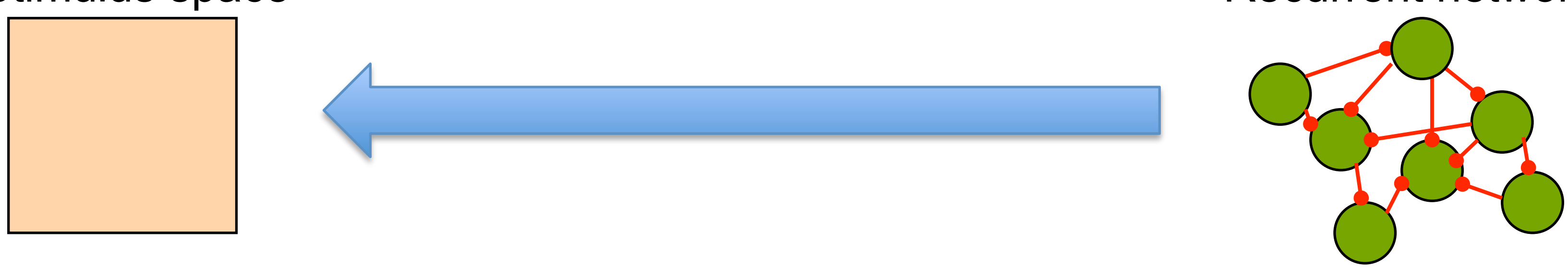
Question: What if the network (i.e. J) is random?

Global topological features (homology groups) of the

Stimulus space

Synaptic weight matrix J

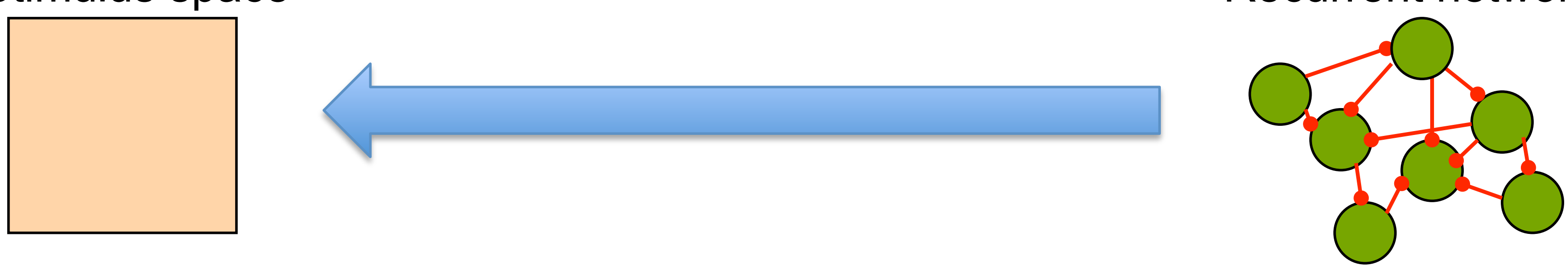
Recurrent network



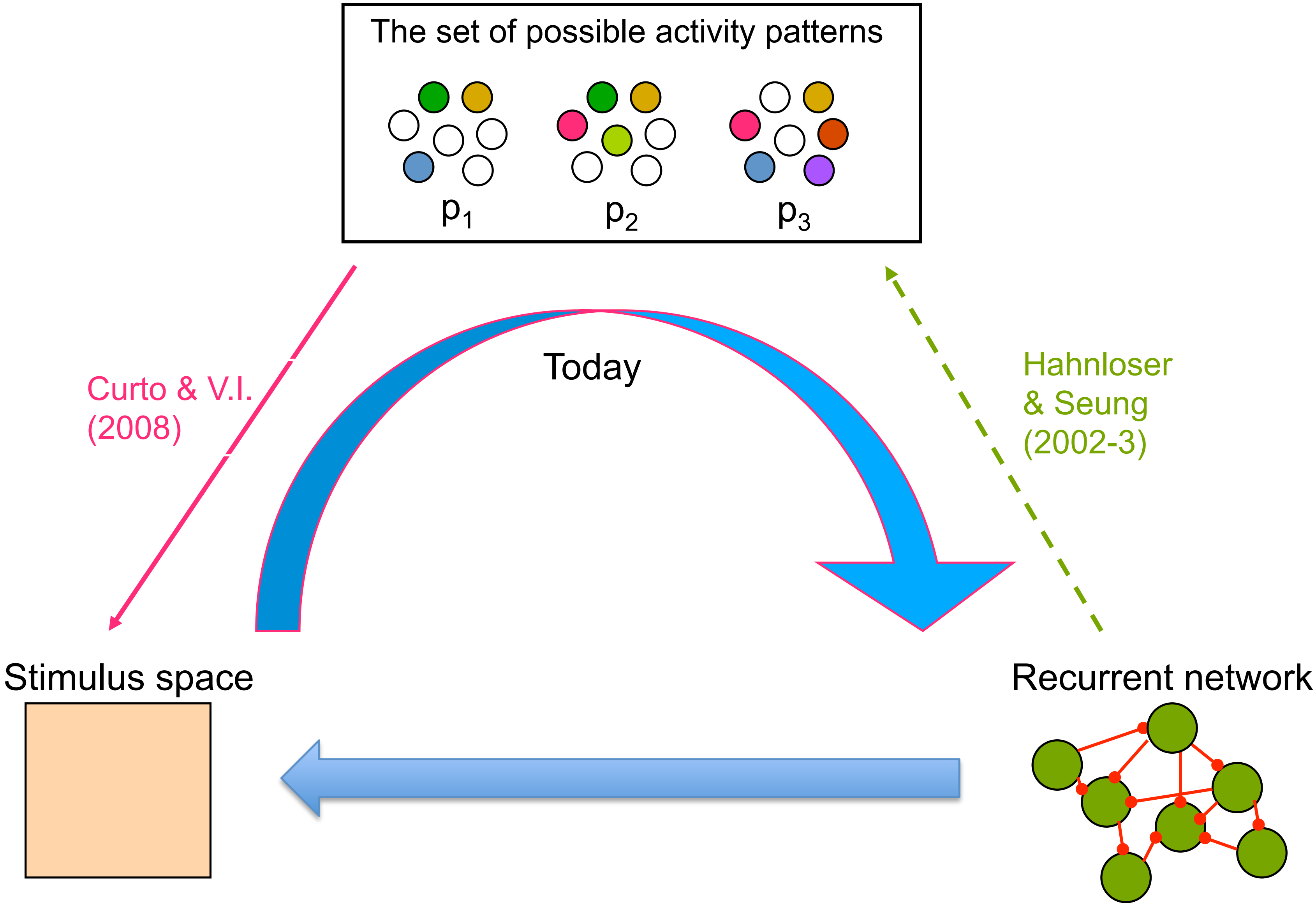
Question: What if the network (i.e. J) is random?

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



Acknowledgements

Supported by the Swartz Foundation.

This is joint work with Carina Curto.

Columbia Center for Theoretical Neuroscience