Flexible networks for memory-encoding

Carina Curto[1], Anda Degeratu[2], Vladimir Itskov[1]

[1]Department of Mathematics, University of Nebraska-Lincoln; [2]Max Planck Institute, Potsdam, Germany.

Introduction

New memories in some brain areas, such as hippocampus, can be encoded quickly. Irrespective of the plasticity mechanism (or learning rule) used to encode memory patterns via changes in synaptic weights, rapid learning is perhaps most easily accomplished if new patterns can be learned via only small modifications of the initial synaptic weights. It may thus be desirable for fast-learning and flexible neural networks to have architectures which enable large numbers of patterns to be encoded by only small perturbations of the synaptic efficacies. What kinds of network architectures have this property?

What kinds of network architectures allow for fast learning of new memory patterns?

Goal: Look for architectures that allow many memories to be encoded via arbitrarily small perturbations of the weights.

$$W_{ij} = J_{ij} + \varepsilon(\Delta W_{ij})$$

The answer should not depend on the particular learning rule used to update the weights.

What exactly do we mean by a memory pattern?

We consider a simple firing rate model:

$$\tau \frac{dx_i}{dt} = -x_i + \varphi\left(\sum_{j=1}^{N} W_{ij} x_j + b_i\right)$$

$x =$ vector of firing rates for each neuron

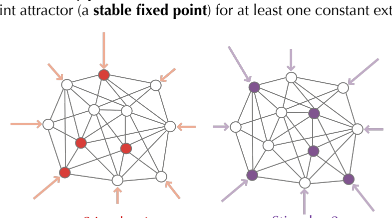
$\tau =$ membrane time constant

$W =$ matrix of synaptic weights

$b =$ vector of external inputs to each neuron

$\varphi =$ nonlinear transfer function

An encoded **memory pattern** of the network **W** is a subset of neurons that co-fires at a fixed point attractor (a **stable fixed point**) for at least one constant external input **b**.



Stimulus 1 Stimulus 2

Different external inputs can activate different memory patterns. The same external input can also activate more than one memory pattern, depending on initial condition.

Wait – why isn't every subset of neurons a memory pattern?

One might expect that, given a subset of neurons, it is always possible to find some constant input **b** for which this subset is exactly the set of neurons that co-fire at a stable fixed point. It would follow that all possible subsets of neurons are memory patterns, irrespective of the network architecture. Fortunately, **this is not true**.

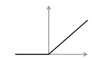
In fact, in this setup the recurrent network acts as a **gating device**. Different memory patterns are activated by different inputs, but some memory patterns can never be activated – even if allowed any possible (constant) feedforward input.

Stable sets in threshold-linear networks

We refer to our encoded memory patterns as **stable sets**.

We also restrict ourselves to the case of **threshold-linear networks**. This just means choosing the nonlinear transfer function to be:

$$\varphi(x) = [x]_+$$



Theorem 1. Stable sets correspond to stable principal submatrices of *-I+W.*

This was previously shown in the case of symmetric matrices in [2], where "stable sets" are called "permitted sets".

original matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

principal submatrix for subset of neurons $\{1,3\}$

$$A_{\{13\}} = \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix}$$

A matrix is *stable* if all its eigenvalues satisfy

$$\mathrm{Re}(\lambda) < 0$$

G-constrained networks (sparse networks)

Not every connection strength in a network is easy to change. A complete lack of anatomical connection is different than a silent synapse, although both cases are represented by 0 in the connectivity matrix.

We say that an architecture matrix J is **constrained by** a simple graph **G** if $J_{ij} = 0$ for all edges $(ij) \notin G$.

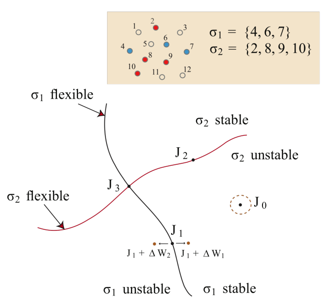
In exploring network flexibility we only allow perturbations of unconstrained entries of the matrix J.

The **clique complex** of G is the set **X(G)** of all cliques in the graph. These are subsets of neurons where all pairwise connections are unconstrained.

Flexible cliques

A *maximally stable clique* of a network W is a stable clique that is not contained in any larger stable clique. A *minimally unstable clique* is an an unstable clique that does not contain any smaller unstable clique.

A **flexible clique** of a network J is a subset of neurons that can become both a maximally stable clique and a minimally unstable clique under arbitrarily small perturbations of the matrix J.



Flexibility of a network J = number of flexible cliques

Question: For a given constraint graph G, how should we choose the weights of the network J such that J has maximal flexibility to learn (and unlearn) new memory patterns?

Structure of maximally flexible networks

Theorem 2. Let G be a graph, X(G) its clique complex, and suppose that $H_1(X(G);Z) = 0$. A G-constrained network J is maximally flexible iff $-I + J$ has a rank 1 completion.

What is a rank 1 completion?

$$-I + J = \begin{pmatrix} -1 & 0 & 2 \\ 0 & -6 & -4 \\ 1 & -3 & -2 \end{pmatrix}$$

A network where (12) edges are constrained to be zero.

$$\begin{pmatrix} -1 & 3 & 2 \\ 2 & -6 & -4 \\ 1 & -3 & -2 \end{pmatrix}$$

A rank 1 completion of –I+J. Only entries constrained to be zero can be filled in with new values.

Why do we need a topological condition?

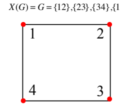
maximally flexible network

$$-I + J = \begin{pmatrix} -1 & 2 & 0 & 1 \\ 1/2 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \end{pmatrix}$$

flexible cliques

$\{12\}$
$\{23\}$
$\{34\}$
$\{14\}$

$X(G) = G = \{12\},\{23\},\{34\},\{14\}$



This network does not have a rank 1 completion!

Good news:

$H_1(X(G);Z) = 0$ is generically satisfied for large random networks that are not overly sparse [3].

For $n = 10^4$ neurons, $H_1(X(G);Z) = 0$ is expected to vanish for connection probability between neurons as low as $p \geq 0.05$.

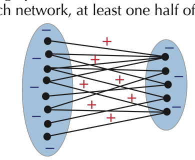
Theorem 3. An unconstrained threshold-linear network J is maximally flexible iff $-I + J$ has rank 1.

Corollary (Thms 2 & 3). All maximally flexible networks are bipartite.

Bipartite networks

A *bipartite network* has a stereotyped sign pattern for the connection strengths between neurons. Note that in any such network, at least one half of the connections must be inhibitory.

$$-I + J = \begin{pmatrix} - & + \\ + & - \end{pmatrix}$$



Conclusions & References

Threshold-linear networks that are maximally flexible have low rank completions and are bipartite. Evidence for such features can be seen even in highly under-sampled connectivity matrices – this gives something new to look for in experiments.

[1] C. Curto, A. Degeratu, V. Itskov. Flexible memory networks. Submitted, 2010. Preprint available at: http://arxiv.org/abs/1009.4958

[2] R.H. Hahnloser, H.S. Seung, J.J. Slotine. Permitted and forbidden sets in symmetric threshold-linear networks. Neural Computation, 2003.

[3] M. Kahle. Topology of random clique complexes. Discrete Math, 2009.