

# Learning and memory with complex synapses

Subhaneil Lahiri and Surya Ganguli

Department of Applied Physics, Stanford University, Stanford CA

## Background

### Storage capacity of synaptic memory

- Capacity limited by new memories overwriting old ones.
- Hopfield capacity  $\propto N$  requires unbounded synaptic strengths.
- Bounded synapses  $\implies$  capacity  $\propto \log N$ .
- Trade-off between learning & remembering.
- Can be ameliorated by using metaplasticity in complex synapses

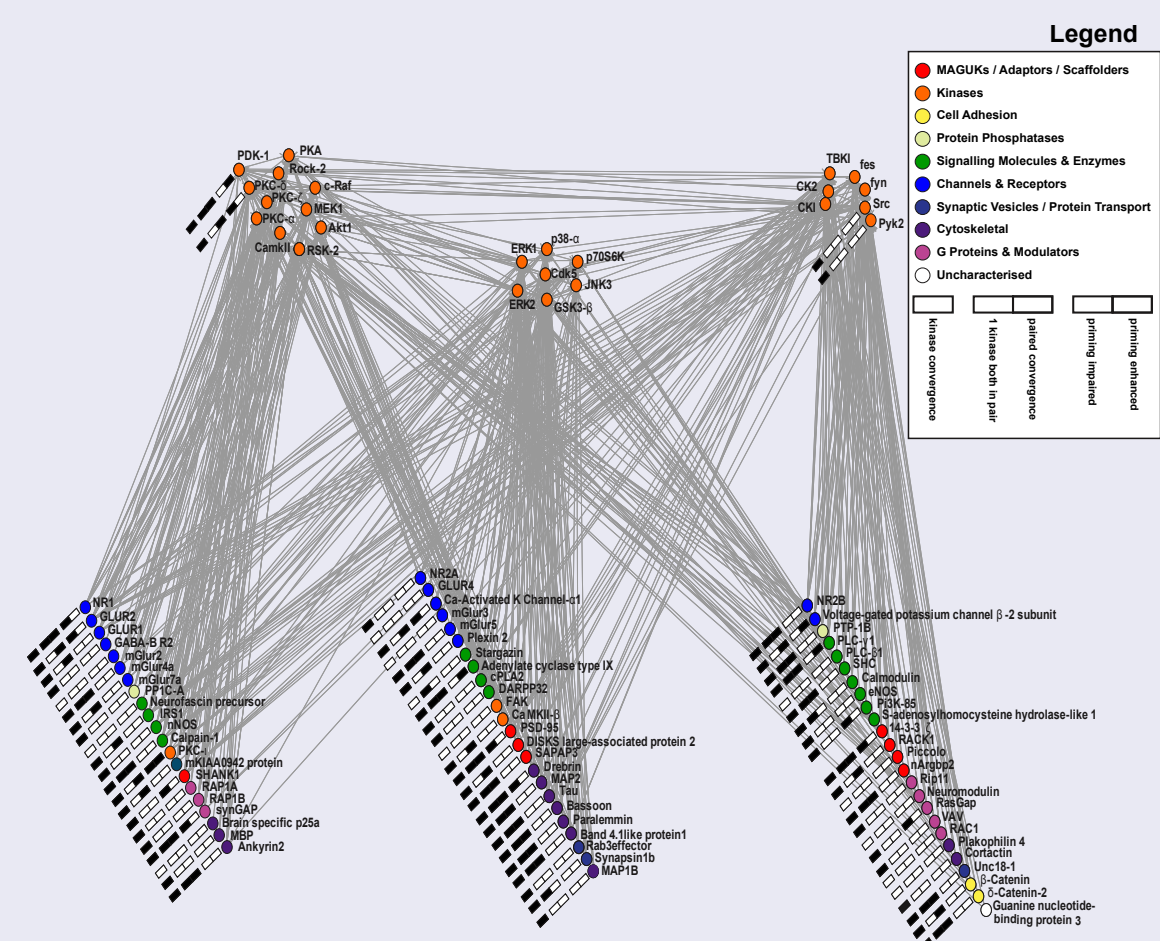
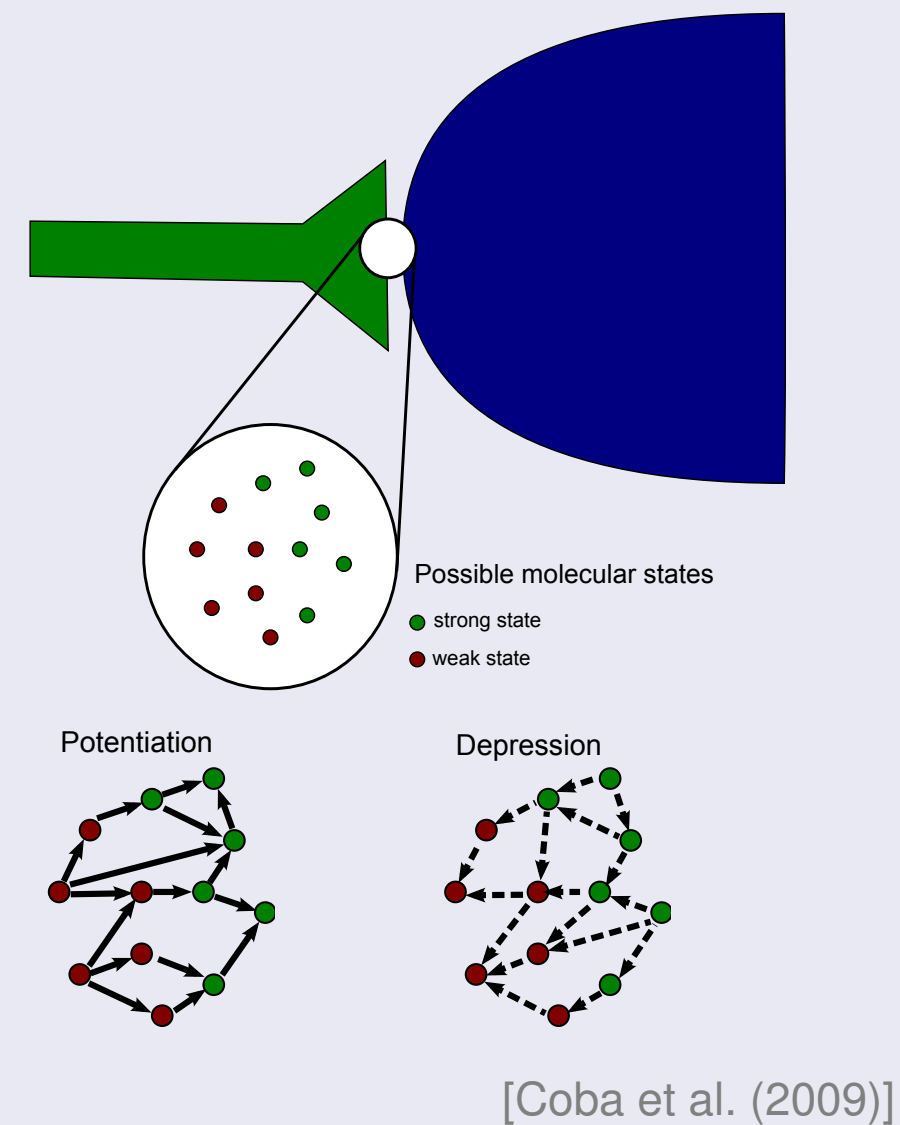
### Complex synapses

Not just a synaptic weight.

Take into account state of molecular network.

Potentiation and depression cause transitions between these states.

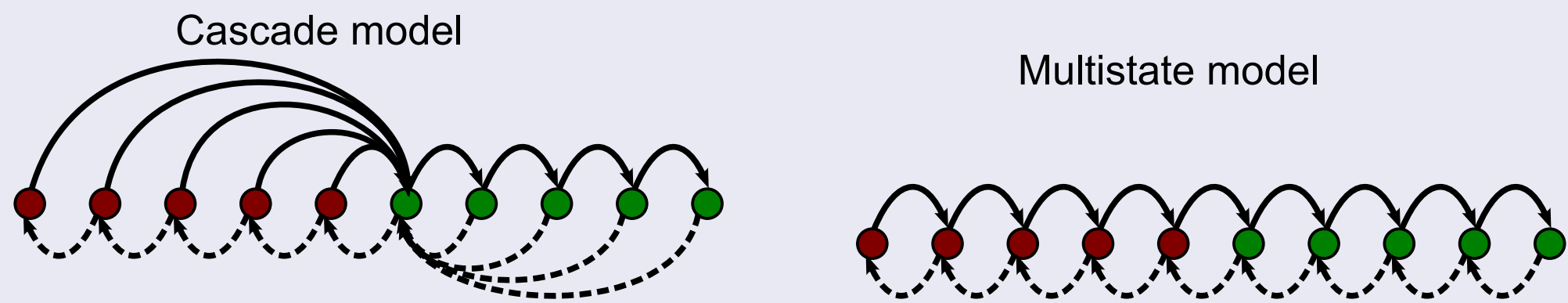
Experimental molecular network



[Coba et al. (2009)]

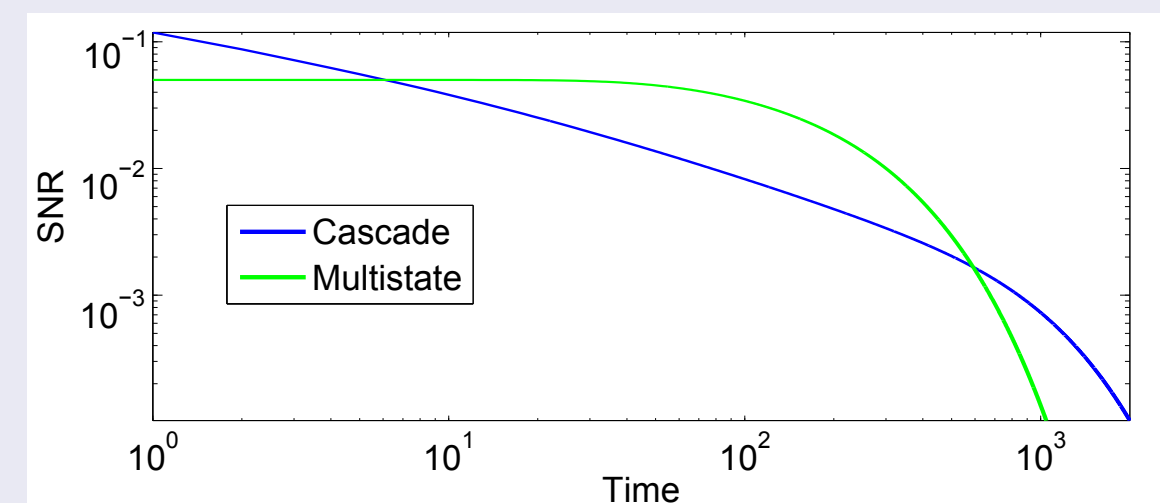
### Cascade and multistate models

Two example models of metaplasticity in complex synapses.



[Fusi et al. (2005), Fusi and Abbott (2007)]

These have different memory storage properties



### Questions

- How is structure of molecular network related to function?
- What are the upper bounds on different measures of memory?
- Which molecular network topologies maximize these measures?

## Framework

### Metaplasticity models

We have  $N$  synapses with  $n$  internal states.

We have two Markov processes describing transition probabilities for potentiation,  $\mathbf{M}^{\text{pot}}$ , and depression,  $\mathbf{M}^{\text{dep}}$ .

Plasticity events are potentiating with probability  $f^{\text{pot}}$  and depressing with probability  $f^{\text{dep}}$ .

After the memory we are tracking, subsequent plasticity events occur at rate  $r$ , with transition probabilities

$$\mathbf{M}^{\text{forget}} = f^{\text{pot}} \mathbf{M}^{\text{pot}} + f^{\text{dep}} \mathbf{M}^{\text{dep}}.$$

This will eventually return it to the equilibrium distribution,  $\mathbf{p}^\infty$ .

### Memory curve

We use the ideal observer approach: read synaptic weights directly. This is an upper bound on what could be read from network activity.

Reconstruction probability of a single synapse:

$$s(t) = f^{\text{pot}} P(\text{strong}, t | \text{pot}, 0) + f^{\text{dep}} P(\text{weak}, t | \text{dep}, 0)$$

Alternatively, if  $\vec{w}$  is an  $N$ -element vector of synaptic strengths,

$$\text{Signal} = \langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) - \vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \rangle$$

$$\text{Noise} = \text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))$$

If we ignore correlations between different synapses, signal-to-noise ratio:

$$\text{SNR}(t) \sim \sqrt{N}(s(t) - s(\infty)).$$

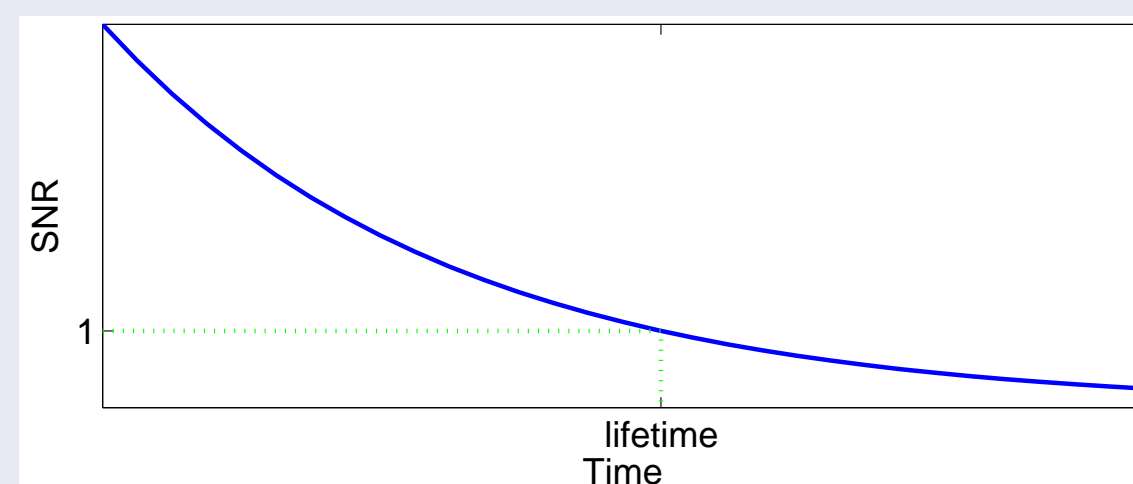
## Upper bounds on performance

### Area bound

We can show that the area under the SNR curve is bounded:

$$A \leq \sqrt{N}(n-1)/r.$$

This leads to a bound on the lifetime of a memory:



$$\text{SNR}(\text{lifetime}) = 1 \implies A \geq \text{lifetime}.$$

This is saturated by a molecular network with the multistate topology.

### Ordering the states

Let  $\mathbf{T}_{ij}$  be the mean first passage time from state  $i$  to state  $j$ . The quantity

$$\eta = \sum_j \mathbf{T}_{ij} \mathbf{p}_j^\infty,$$

is independent of the initial state  $i$ . It is known as Kemeny's constant. [Kemeny and Snell (1960)]

We define:

$$\eta_i^+ = \sum_{j \in \text{strong}} \mathbf{T}_{ij} \mathbf{p}_j^\infty, \quad \eta_i^- = \sum_{j \in \text{weak}} \mathbf{T}_{ij} \mathbf{p}_j^\infty.$$

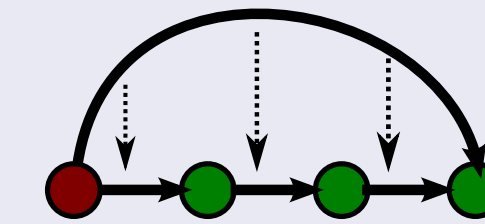
These measure "distance" to the strong/weak states. They can be used to put the states in order (increasing  $\eta^-$  or decreasing  $\eta^+$ ).

## Maximal area

Given any molecular network, we can construct one with the multistate topology that has

- same state order,
- same equilibrium distribution,
- larger area.

Uses a deformation that reduces "shortcut" transition probabilities and increases the bypassed "direct" ones.



The area of this model is

$$A = \frac{2\sqrt{N}}{r} \sum_k \mathbf{p}_k^\infty |k - \langle k \rangle|.$$

Maximum is when all probability is at ends.

## Envelope memory curve

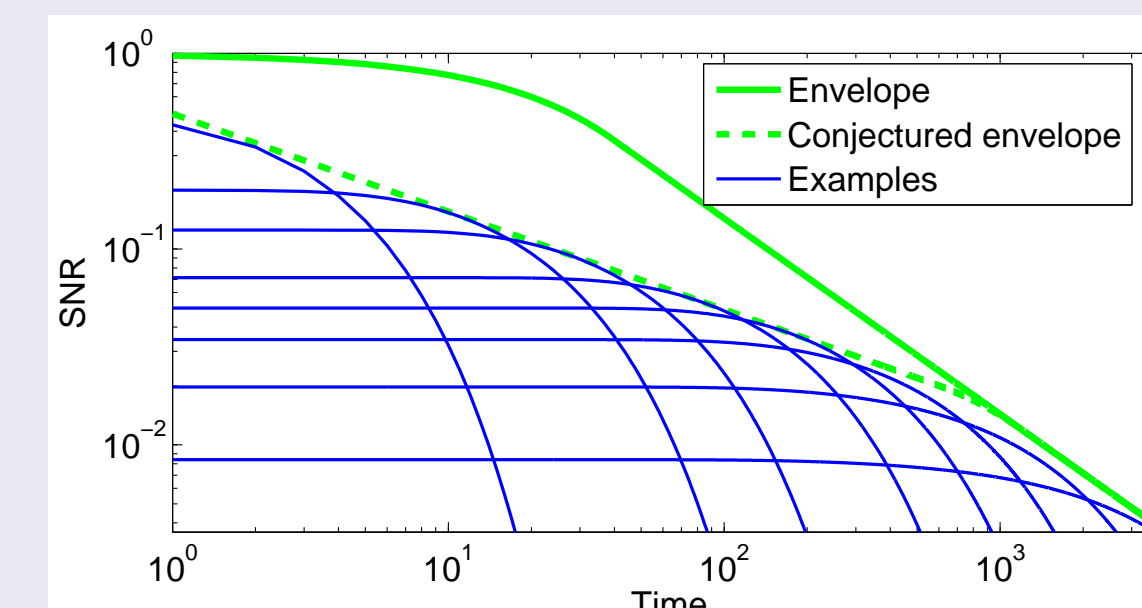
### Maximal SNR curve

Markov process  $\implies$  SNR is sum of exponentials.

If we find the maximum sum of exponentials **at one time** subject to upper bounds on initial SNR and area, we get an upper bound on SNR at that time.

Resulting curve is always a single exponential.

If we vary the time at which we find the optimum, we get an envelope curve with a power law tail.



## Extra constraint

The envelope above may not be tight.

We can get a tight envelope – one that can be saturated at any single time by some model – if we add one more constraint.

Schematically, mode by mode:

$$\text{SNR}(0) \sqrt{\text{time-scale}} \leq \sqrt{N} \cdot \mathcal{O}(1).$$

We have found no model can exceed this. It is saturated by a diffusive chain:

$$\text{SNR}(0) \sim \frac{1}{n}, \quad \text{time-scale} \sim n^2.$$

## Maximum lifetime

We can use the envelope to get a stricter bound on the lifetime of a memory

$$\text{Envelope}(\text{max lifetime}) = 1, \quad \text{max lifetime} = \frac{\sqrt{N}(n-1)}{er}.$$

## References

- M. P. Coba, A. J. Pocklington, M. O. Collins, M. V. Kopanitsa, R. T. Uren, S. Swamy, M. D. Croning, J. S. Choudhary, and S. G. Grant, "Neurotransmitters drive combinatorial multistate postsynaptic density networks", *Sci Signal*, 2(68):ra19, 2009, PubMed: 19401593, [PubMed Central:PMC3280897].
- S. Fusi, P. J. Drew, and L. F. Abbott, "Cascade models of synaptically stored memories", *Neuron*, 45(4):599–611, Feb 2005, PubMed: 15721245.
- S. Fusi and L. F. Abbott, "Limits on the memory storage capacity of bounded synapses", *Nat. Neurosci.*, 10(4): 485–493, Apr 2007, PubMed: 17351638.
- J.G. Kemeny and J.L. Snell, *Finite markov chains*. Springer, 1960.