Learning and memory with complex synapses

Subhaneil Lahiri and Surya Ganguli
Department of Applied Physics, Stanford University, Stanford CA

Background

Storage capacity ofsynaptic memory

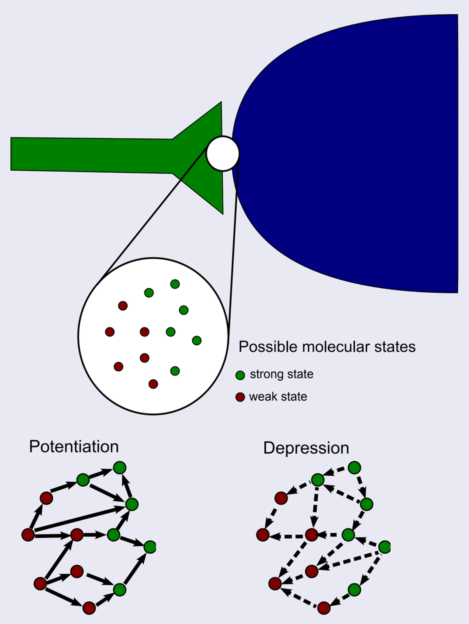
- Capacity limited by new memories overwriting old ones.
- Hopfield capacity $\propto N$ requires unbounded synaptic strengths.
- Bounded synapses $\implies$ capacity $\propto \log N$.
- Trade-off between learning & remembering.
- Can be ameliorated by using metaplasticity in complex synapses

Complex synapses

Not just a synaptic weight.

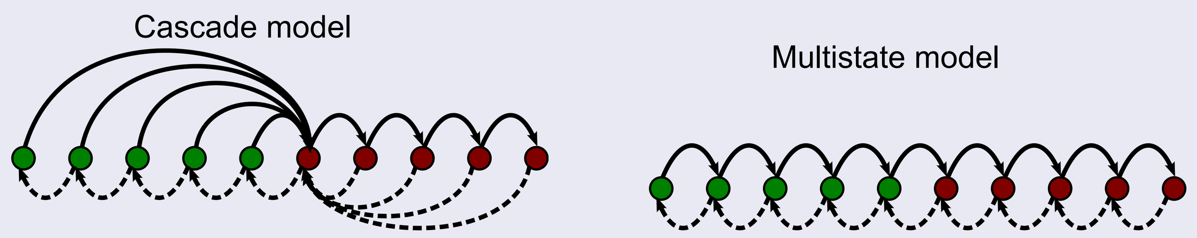
Take into account state of molecular network.

Potentiation and depression cause transitions between these states.

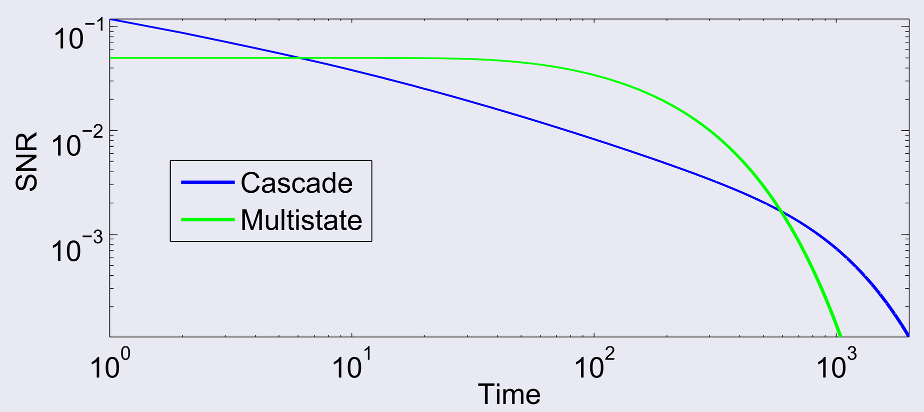


Cascade and multistate models

Two models of metaplasticity in complex synapses.



These have different memory storage properties



Framework

Markov processes

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^{pot} and depressing with probability f^{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 $\mathbf{W}$ is an N-element vector of synaptic strengths,

$$\text{Signal} = \langle \mathbf{W}_{\text{ideal}} \cdot \mathbf{W}(t) - \mathbf{W}_{\text{ideal}} \cdot \mathbf{W}(\infty) \rangle$$
$$\text{Noise} = \text{Var}\left(\mathbf{W}_{\text{ideal}} \cdot \mathbf{W}(\infty)\right)$$

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:

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

This is saturated by a molecular networkwith the multistate topology.

Ordering the states

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

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

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

We define:

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

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

Maximal area

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

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

Maximum is when all probability is at ends.

Envelope memory curve

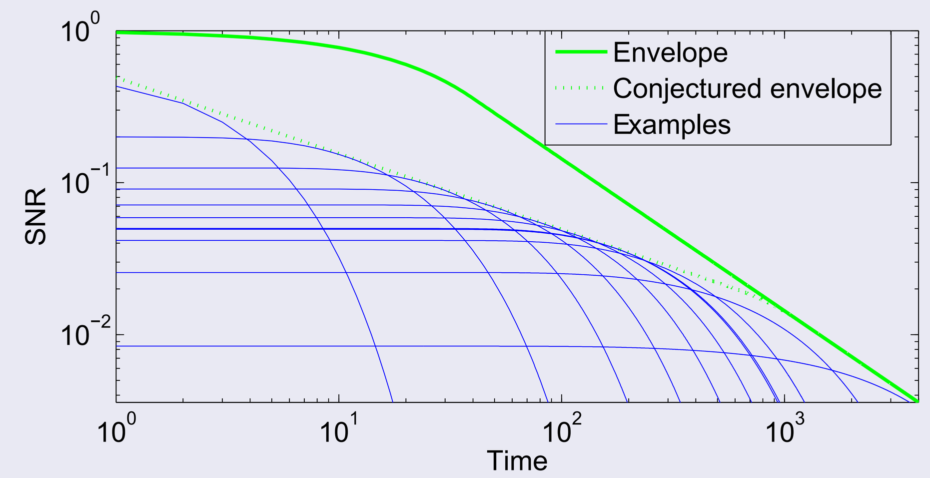
Maximal SNR curve

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

If we find the optimum 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 isn't 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:

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

We have found no model that exceeds this.

Maximum lifetime

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

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

References

J.G. Kemeny and J.L. Snell, *Finite markov chains*. Springer, 1960.