A general theory of learning and memory with Complex Synapses

based on work with Surya Ganguli

Subhaneil Lahiri

Stanford University, Applied Physics

April 9, 2013

Complex synapses

2013-04-09

A general theory of learning and memory with Complex Synapses
based on work with Surya Ganguli
Subhaneil Lahiri
Stanford University, Applied Physics
April 9, 2013

Introduction

We often model synaptic plasticity as the change of a single number (synaptic weight). In reality, there is a complex dynamical system inside a synapse.

Semi-realistic models of synaptic memory have terrible storage without synaptic complexity.

We will study the entire space of a broad class of models of complex synapses to find upper bounds on their performance.

Introduction
We often model synaptic plasticity as the change of a single number (synaptic weight). In reality, there is a complex dynamical system inside a synapse.
Semi-realistic models of synaptic memory have terrible storage without synaptic complexity.
We will study the entire space of a broad class of models of complex synapses to find upper bounds on their performance.

1. amplitude of psp.
2. finite number of values.

Outline

1. Why complex synapses?
2. Modelling synaptic complexity
3. Upper bounds
4. Envelope memory curve

Outline
1. Why complex synapses?
2. Modelling synaptic complexity
3. Upper bounds
4. Envelope memory curve

1. review terrible properties of simple synapses.
2. mathematical formalism of model, quantify performance (memory decay over time)
3. upper bounds on single numbers that describe performance at all times
4. upper bounds at finite times

Section 1

Why complex synapses?

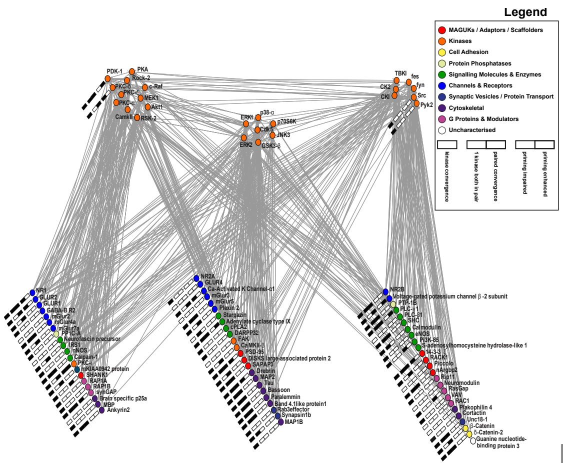
2013-04-09

Complex synapses
└─Why complex synapses?

Section 1

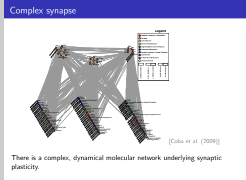
Why complex synapses?

Complex synapse



[Coba et al. (2009)]

There is a complex, dynamical molecular network underlying synaptic plasticity.



1. Does this matter?
2. Could just be the machinery for changing synaptic weight
3. link back to questions on "There"

Storage capacity of synaptic memory

A classical perceptron (used as a recognition memory device) has a capacity $\propto N$, the number of synapses.

Requires synapses' dynamic range also $\propto N$.

If we restrict synaptic weight to a fixed, finite set of values,
$\implies$ tradeoff between learning and forgetting:
new memories overwriting old.

If we wish to store new memories rapidly, memory capacity $\sim \mathcal{O}(\log N)$.
[Amit and Fusi (1992), Amit and Fusi (1994)]

To circumvent this tradeoff, need to go beyond model of a synapse as a single number.

Storage capacity of synaptic memory

A classical perceptron (used as a recognition memory device) has a capacity $\propto N$, the number of synapses.

Requires synapses' dynamic range also $\propto N$.

If we restrict synaptic weight to a fixed, finite set of values,
$\implies$ tradeoff between learning and forgetting:
new memories overwriting old.

If we wish to store new memories rapidly, memory capacity $\sim \mathcal{O}(\log N)$.
[Amit and Fusi (1992), Amit and Fusi (1994)]

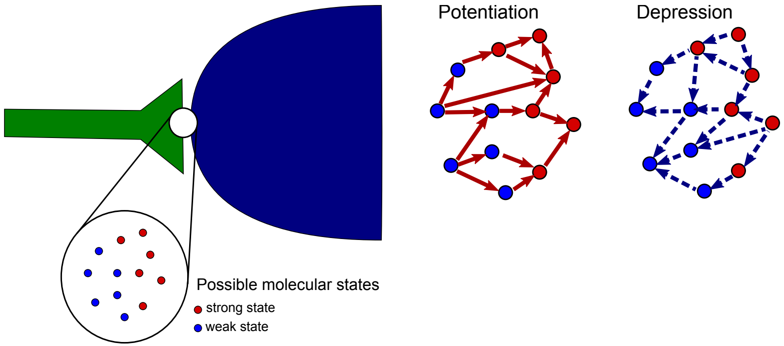
To circumvent this tradeoff, need to go beyond model of a synapse as a single number.

Section 2

Modelling synaptic complexity

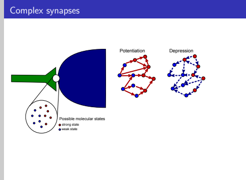
Section 2
Modelling synaptic complexity

Complex synapses



2013-04-09

Complex synapses
└Modelling synaptic complexity
└Complex synapses



Simplifying assumptions

- There are N identical synapses with M internal functional states.
- States of different synapses are independent of each other.
- Which synapses eligible for plasticity chosen randomly.
- Potentiating/depressing plasticity events $\sim$ Poisson processes with rates $rf^{\text{pot/dep}}$, where $f^{\text{pot}} + f^{\text{dep}} = 1$.
- Potentiation and depression are described by Markov processes with transition probabilities $\mathbf{M}^{\text{pot/dep}}$.
- Synaptic weights of the internal states are given by vector $\mathbf{w}$. Can only take values ± 1.

Simplifying assumptions

- There are N identical synapses with M internal functional states.
- States of different synapses are independent of each other.
- Which synapses eligible for plasticity chosen randomly.
- Potentiating/depressing plasticity events $\sim$ Poisson processes with rates $rf^{\text{pot/dep}}$, where $f^{\text{pot}} + f^{\text{dep}} = 1$.
- Potentiation and depression are described by Markov processes with transition probabilities $\mathbf{M}^{\text{pot/dep}}$.
- Synaptic weights of the internal states are given by vector $\mathbf{w}$. Can only take values ± 1.

In other words, r is the total rate of plasticity events per synapse and $f^{\text{pot/dep}}$ are the fraction of these events that are potentiating/depressing.

Probabilities

At $t = 0$, the memory is created by $\mathbf{M}^{\text{pot/dep}}$ with probability $f^{\text{pot/dep}}$.

Forgetting caused by subsequent memories, evolving as

$$\frac{\mathrm{d}\mathbf{p}(t)}{\mathrm{d}t} = r\mathbf{p}(t)\mathbf{W}^{\mathrm{F}}, \qquad \mathbf{W}^{\mathrm{F}} = f^{\text{pot}}\mathbf{M}^{\text{pot}} + f^{\text{dep}}\mathbf{M}^{\text{dep}} - \mathbf{I},$$

Eventually, this will settle into the equilibrium distribution:

$$\mathbf{p}^{\infty}\mathbf{W}^{\mathrm{F}} = 0.$$

Probabilities

At $t = 0$, the memory is created by $\mathbf{M}^{\text{pot/dep}}$ with probability $f^{\text{pot/dep}}$.

Forgetting caused by subsequent memories, evolving as

$$\frac{\mathrm{d}\mathbf{p}(t)}{\mathrm{d}t} = r\mathbf{p}(t)\mathbf{W}^{\mathrm{F}}, \qquad \mathbf{W}^{\mathrm{F}} = f^{\text{pot}}\mathbf{M}^{\text{pot}} + f^{\text{dep}}\mathbf{M}^{\text{dep}} - \mathbf{I},$$

Eventually, this will settle into the equilibrium distribution:

$$\mathbf{p}^{\infty}\mathbf{W}^{\mathrm{F}} = 0.$$

1. for this one, we keep track of pot/dep
2. $\mathbf{W}^{\mathrm{F}}$ is forgetting matrix, don't keep track of pot/dep
3. In equilibrium prior to memory creation

Memory curve

$\vec{w}$ is the N-element vector of synaptic weights.

$$\begin{aligned}\text{Signal} &= \langle \vec{w}_{\text{ideal}} \cdot \vec{w}(t) - \vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \rangle \\ \text{Noise} &= \text{Var}\left(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty) \right)\end{aligned}$$

Related to reconstruction probability of single synapses.

$$\text{SNR}(t) \sim \sqrt{N}\, P(\text{strong/weak}, t | \text{pot/dep}, t=0) - \ldots (t=\infty).$$

Memory curve

$\vec{w}$ is the N-element vector of synaptic weights.

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

Related to reconstruction probability of single synapses.

$\text{SNR}(t) \sim \sqrt{N}\, P(\text{strong/weak}, t | \text{pot/dep}, t=0) - \ldots (t=\infty).$

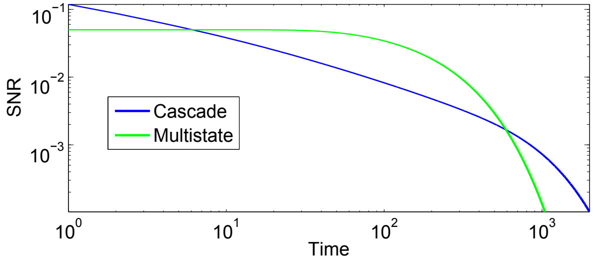
1. of different synapses
2. ideal: pot→strong...
3. if we ignore correlations...

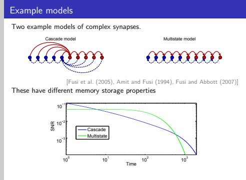
Example models

Two example models of complex synapses.

[Fusi et al. (2005), Amit and Fusi (1994), Fusi and Abbott (2007)]

These have different memory storage properties





Questions

- Can we understand the space of *all possible* synaptic models?
- How does structure (topology) of model $\rightarrow$ function (memory curve)?
- What are the limits on what can be achieved?
- Which transition topologies saturate these limits?

Questions

- Can we understand the space of *all possible* synaptic models?
- How does structure (topology) of model $\rightarrow$ function (memory curve)?
- What are the limits on what can be achieved?
- Which transition topologies saturate these limits?

1. not just individual models
2. understand net (link on topology)
3. avoid using word "optimal". depends on what want to do.

Memory curve 2

Memory curve given by

$$\mathrm{SNR}(t) = \frac{\sqrt{N}(2f^{\text{pot}}f^{\text{dep}})}{\sqrt{\mathbf{p}^{\infty}_{+}\mathbf{p}^{\infty}_{-}}}\,\mathbf{p}^{\infty}\left(\mathbf{M}^{\text{pot}} - \mathbf{M}^{\text{dep}}\right)\exp\left(rt\mathbf{W}^{\mathrm{F}}\right)\mathbf{w}.$$

Constraints: $\mathbf{M}^{\text{pot/dep}}_{ij} \in [0,1], \qquad \sum_j \mathbf{M}^{\text{pot/dep}}_{ij} = 1.$

Eigenmode decomposition:

$$\mathrm{SNR}(t) = \sqrt{N}\sum_a \mathcal{I}_a\, \mathrm{e}^{-rt/\tau_a}.$$

1. prefactors don't do anything, ignore
2. prior state, encoding, forgetting, readout
3. difficult to to apply
4. what are constraints on these?

Section 3

Upper bounds

Section 3
Upper bounds

Initial SNR as flux

Initial SNR is closely related to flux between strong & weak states

$$\mathrm{SNR}(0) \leq \frac{4\sqrt{N}}{r}\,\boldsymbol{\Phi}_{-+}.$$

Max when potentiation guarantees $\mathbf{w} \to +1$,
depression guarantees $\mathbf{w} \to -1$.

Initial SNR as flux
Initial SNR is closely related to flux between strong & weak states
$\mathrm{SNR}(0) \leq \frac{4\sqrt{N}}{r}\boldsymbol{\Phi}_{-+}.$
Max when potentiation guarantees $\mathbf{w} \to +1$,
depression guarantees $\mathbf{w} \to -1$.

1. usually saturated: pot never dec, dep never inc

Section 4

Envelope memory curve

Section 4
Envelope memory curve

Acknowledgements

Thanks to:

- Surya Ganguli
- Stefano Fusi
- Marcus Benna

2013-04-09

Complex synapses
└─Envelope memory curve
└─Acknowledgements

Acknowledgements

Thanks to:

- Surya Ganguli
- Stefano Fusi
- Marcus Benna

1. Last slide!

References I

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.
6

D. J. Amit and S. Fusi.
"Constraints on learning in dynamic synapses".
Network: Computation in Neural Systems, 3(4):443–464, 1992.
7

D. J. Amit and S. Fusi.
"Learning in neural networks with material synapses".
Neural Computation, 6(5):957–982, 1994.
7 13

References I
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.
D. J. Amit and S. Fusi.
"Constraints on learning in dynamic synapses".
Network: Computation in Neural Systems, 3(4):443–464, 1992.
D. J. Amit and S. Fusi.
"Learning in neural networks with material synapses".
Neural Computation, 6(5):957–982, 1994.

References II

S. Fusi, P. J. Drew, and L. F. Abbott.
"Cascade models of synaptically stored memories".
Neuron, 45(4):599–611, Feb 2005, PubMed:15721245.
13

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.
13

References II

S. Fusi, P. J. Drew, and L. F. Abbott.
"Cascade models of synaptically stored memories".
Neuron, 45(4):599–611, Feb 2005, PubMed:15721245.
13

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.
13