Learning and memory with complex synaptic plasticity

Subhaneil Lahiri and Surya Ganguli

Stanford University, Applied Physics

February 11, 2016

Introduction

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

Discrete models of synaptic plasticity have terrible memory 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.

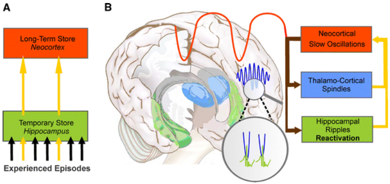
This leads to understanding of what structures are useful for storing memories for different timescales.

Timescales of memory

Memories stored in different places for different timescales

[Squire and Alvarez (1995)]

[McClelland et al. (1995)]



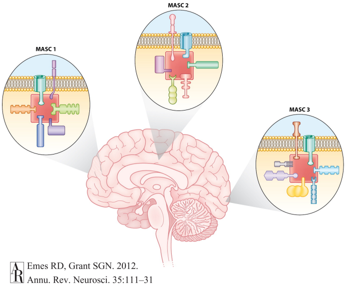
[Born and Wilhelm (2012)]

Also: Cerebellar cortex → nuclei.

[Attwell et al. (2002)]

[Cooke et al. (2004)]

Different synapses have different molecular structures.



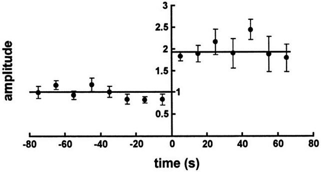
[Emes and Grant (2012)]

Storage capacity of synaptic memory

A classical perceptron has a capacity $\propto N$, ($\#$ synapses).

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

With discrete, finite synapses:
$\Longrightarrow$ new memories overwrite old,
$\Longrightarrow$ stability-plasticity dilemma.



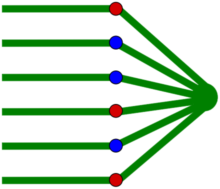
[Petersen et al. (1998), O'Connor et al. (2005)]

When we store new memories rapidly, memory capacity $\sim \mathcal{O}(\log N)$.

[Amit and Fusi (1992), Amit and Fusi (1994)]

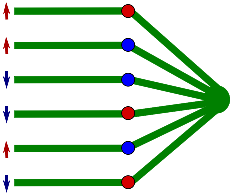
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



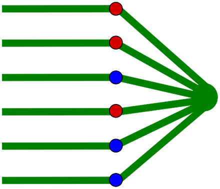
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



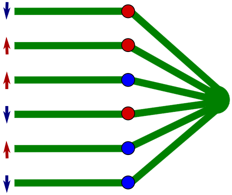
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



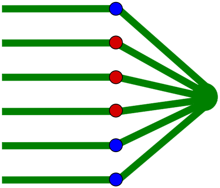
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



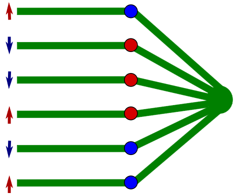
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



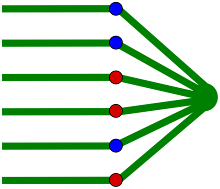
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



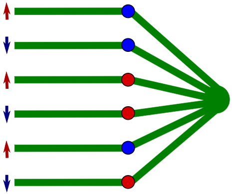
Recognition memory

Synapses given a sequence of patterns (pot & dep) to store



Recognition memory

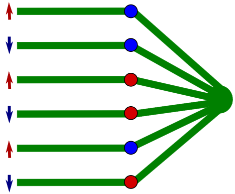
Synapses given a sequence of patterns (pot & dep) to store



Later: presented with a pattern. Has it been seen before?

Recognition memory

Synapses given a sequence of patterns (pot & dep) to store

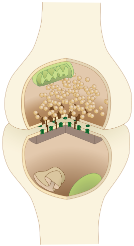


Later: presented with a pattern. Has it been seen before?

Compare $\vec{s}\cdot\vec{w}(t)$ to threshold. [Sommer and Dayan (1998)]

$$\mathrm{SNR}(t) = \frac{\langle \vec{s}\cdot\vec{w}(t)\rangle - \langle \vec{s}\cdot\vec{w}(\infty)\rangle}{\sqrt{\mathrm{Var}(\vec{s}\cdot\vec{w}(\infty))}}, \qquad \overline{\mathrm{SNR}}(\tau) = \int \mathrm{d}\tau\, \frac{\mathrm{e}^{-t/\tau}}{\tau}\,\mathrm{SNR}(t).$$

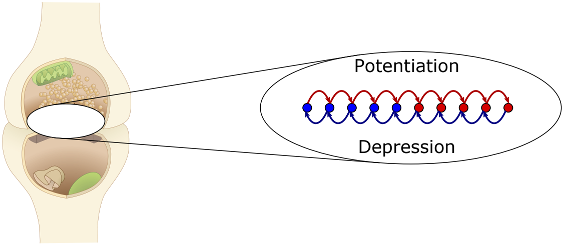
Models of complex synaptic dynamics



Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

● weak
● strong

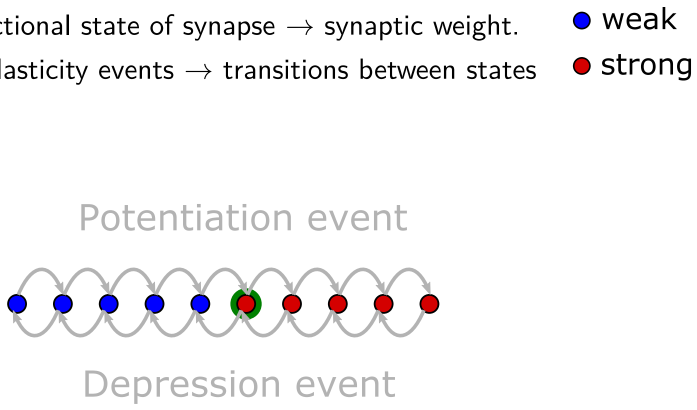


States: #AMPAR, #NMDAR, NMDAR subunit composition,
CaMK II autophosphorylation, activating PKC, p38 MAPK,...

[Fusi et al. (2005), Fusi and Abbott (2007), Barrett and van Rossum (2008)]

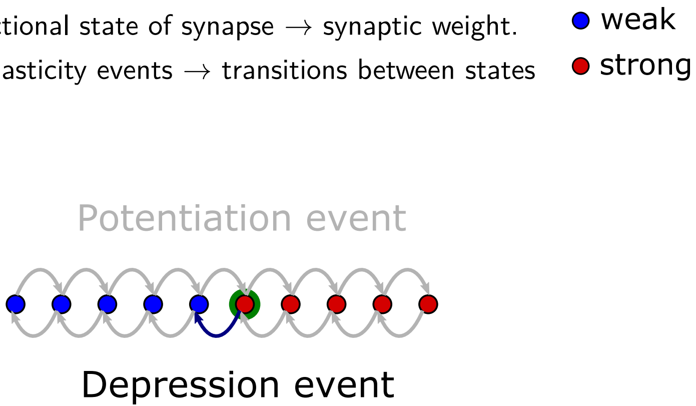
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



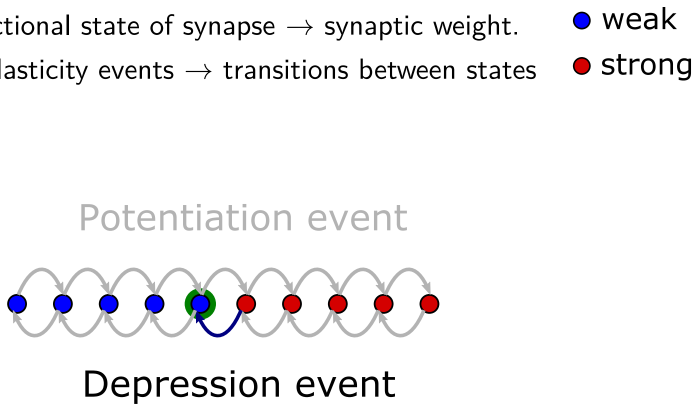
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



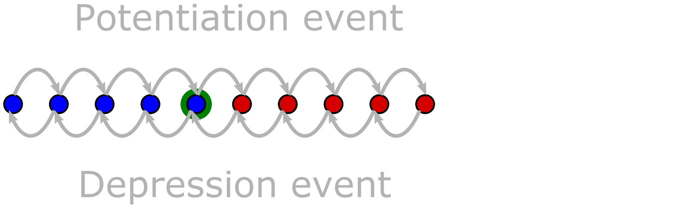
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



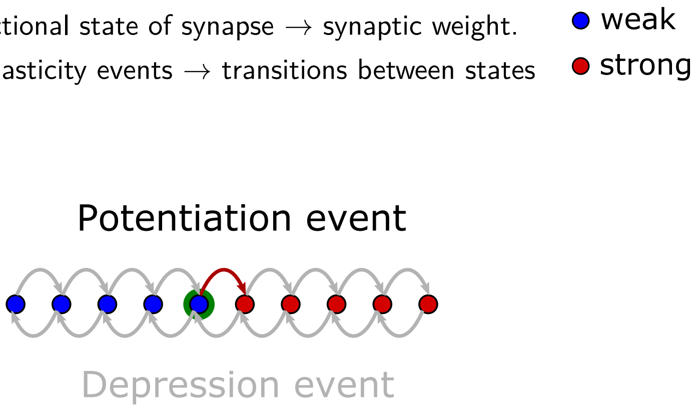
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



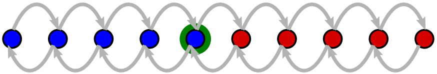
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



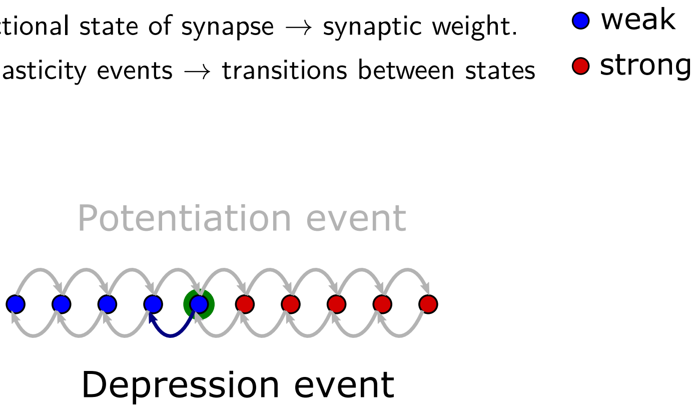
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

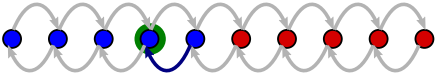
Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

weak

strong

Potentiation event

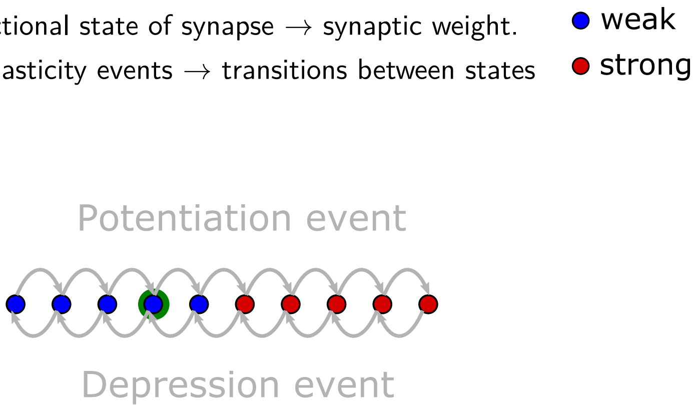


Depression event

Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



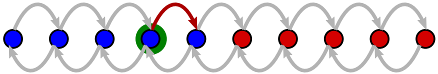
Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

weak
strong

Potentiation event

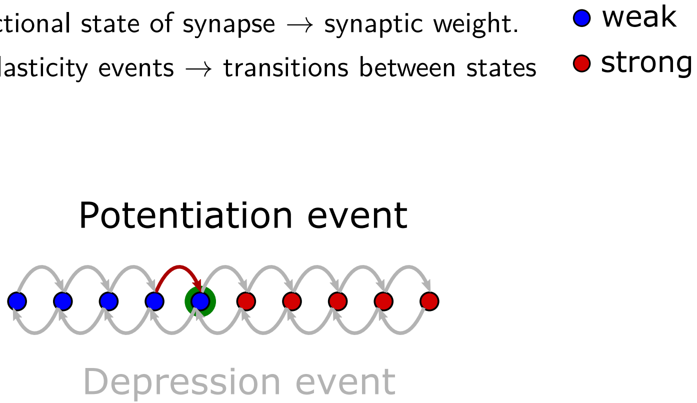


Depression event

Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

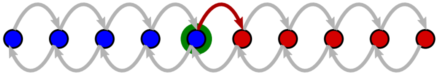


Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

Potentiation event

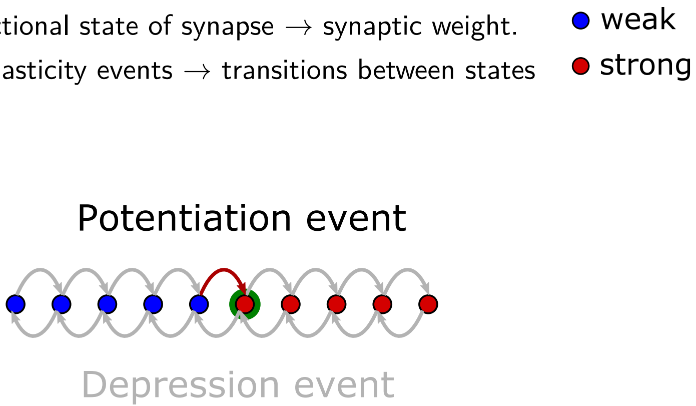


Depression event

Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



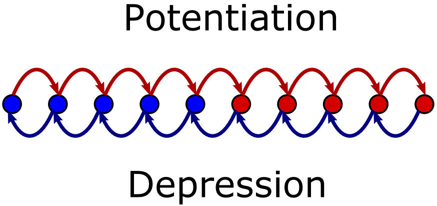
Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

Models of complex synaptic dynamics

- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

weak

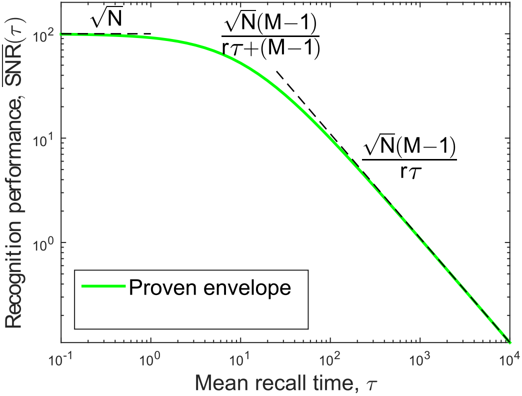
strong



Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

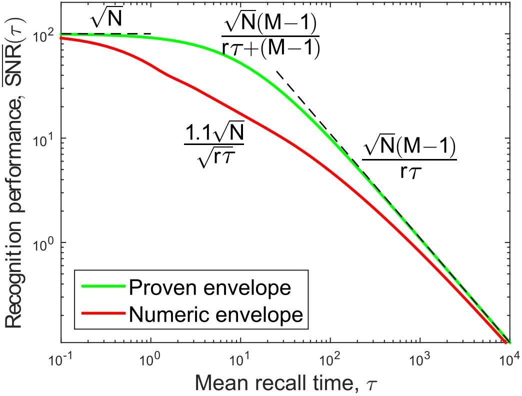
Proven envelope: memory frontier

Upper bound on memory curve at *any* time.

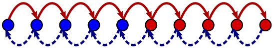


Proven envelope: memory frontier

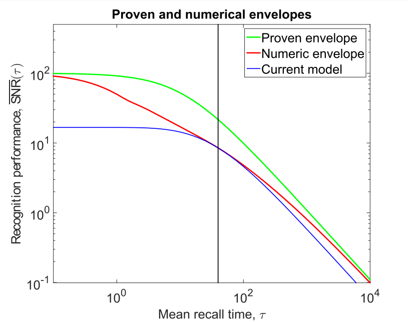
Upper bound on memory curve at *any* time.

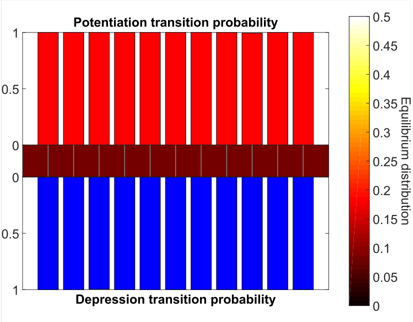


Serial topology:

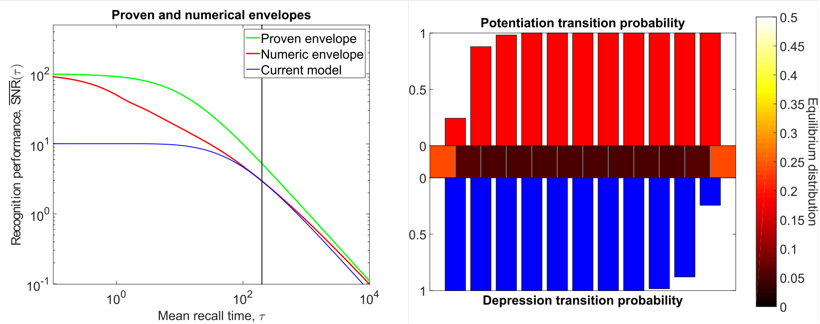


Models that maximise memory for one timescale

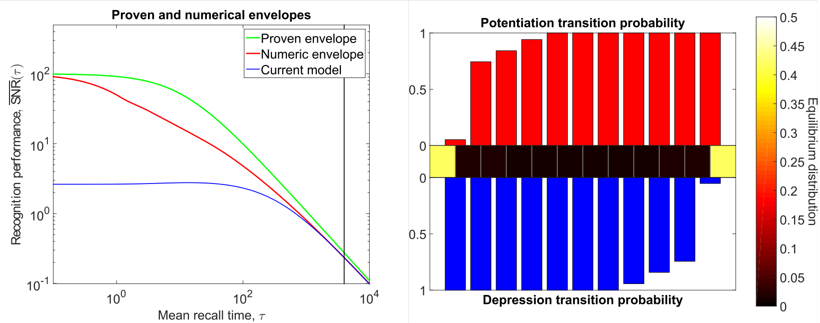




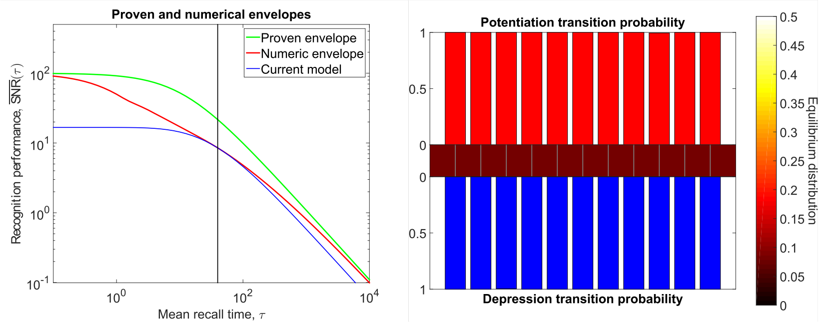
Models that maximise memory for one timescale



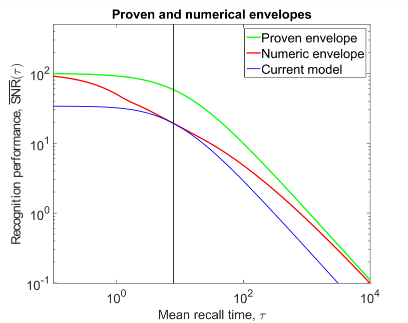
Models that maximise memory for one timescale

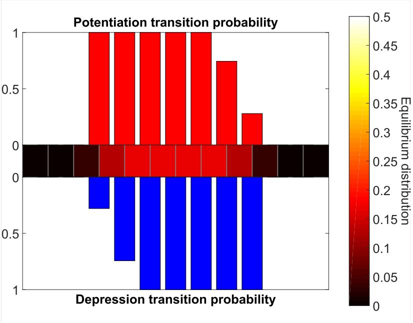


Models that maximise memory for one timescale

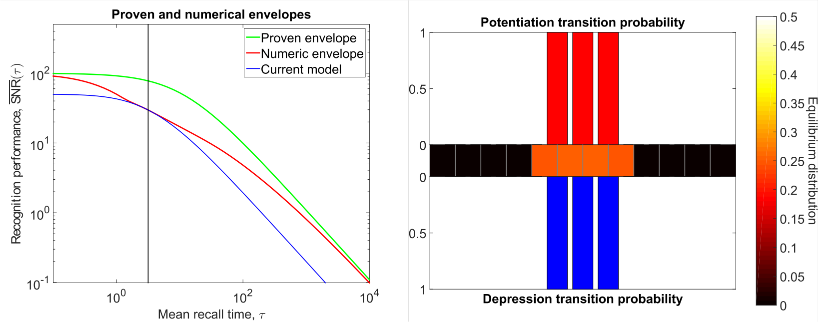


Models that maximise memory for one timescale

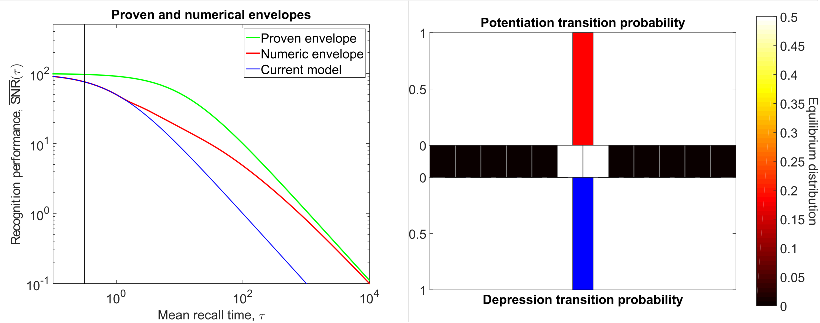




Models that maximise memory for one timescale



Models that maximise memory for one timescale



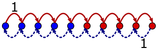
Synaptic structures for different timescales of memory

Real synapses limited by molecular building blocks.
Evolution had larger set of priorities.

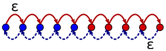
What can we conclude?

Short timescales $\longrightarrow$ Intermediate timescales $\longrightarrow$ Long timescales

$\longrightarrow$



$\longrightarrow$



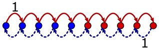
Synaptic structures for different timescales of memory

Real synapses limited by molecular building blocks.
Evolution had larger set of priorities.

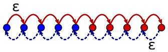
What can we conclude?

Short timescales $\longrightarrow$ Intermediate timescales $\longrightarrow$ Long timescales

$\longrightarrow$



$\longrightarrow$

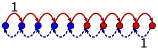
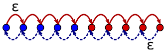


short & wide $\longrightarrow$ long & thin

Synaptic structures for different timescales of memory

Real synapses limited by molecular building blocks.
Evolution had larger set of priorities.

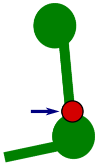
What can we conclude?

Short timescales	$\longrightarrow$	Intermediate timescales	$\longrightarrow$	Long timescales
	$\longrightarrow$		$\longrightarrow$	
short & wide	$\longrightarrow$	long & thin		
		strong transitions	$\longrightarrow$	weak transitions

Proposed Experimental design

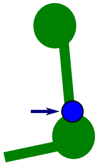
Subject a synapse to a sequence of candidate plasticity events.

Observe the changes in synaptic efficacy.



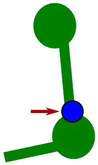
Proposed Experimental design

Subject a synapse to a sequence of candidate plasticity events.
Observe the changes in synaptic efficacy.



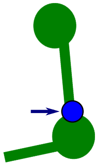
Proposed Experimental design

Subject a synapse to a sequence of candidate plasticity events.
Observe the changes in synaptic efficacy.



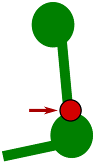
Proposed Experimental design

Subject a synapse to a sequence of candidate plasticity events.
Observe the changes in synaptic efficacy.



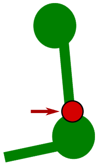
Proposed Experimental design

Subject a synapse to a sequence of candidate plasticity events.
Observe the changes in synaptic efficacy.



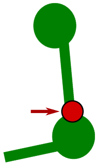
Proposed Experimental design

Subject a synapse to a sequence of candidate plasticity events.
Observe the changes in synaptic efficacy.



Proposed Experimental design

Subject a synapse to a sequence of candidate plasticity events.
Observe the changes in synaptic efficacy.

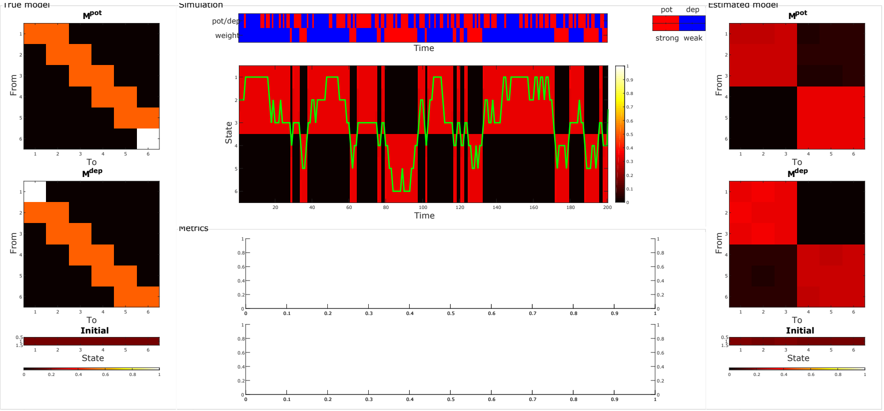


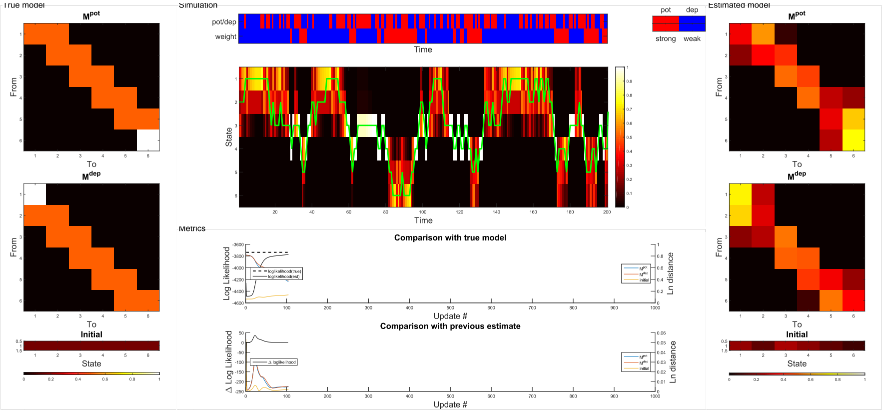
EM algorithms:

Sequence of hidden states $\rightarrow$ estimate transition probabilities
Transition probabilities $\rightarrow$ estimate sequence of hidden states

[Baum et al. (1970), Rabiner and Juang (1993), Dempster et al. (2007)]

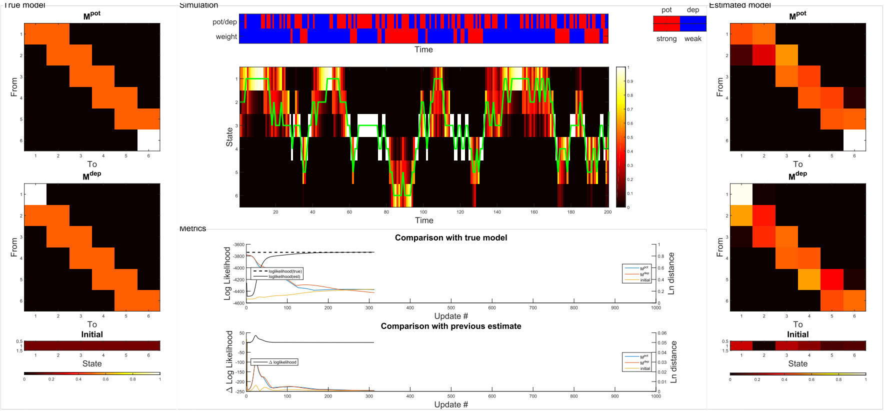
Fitting algorithm



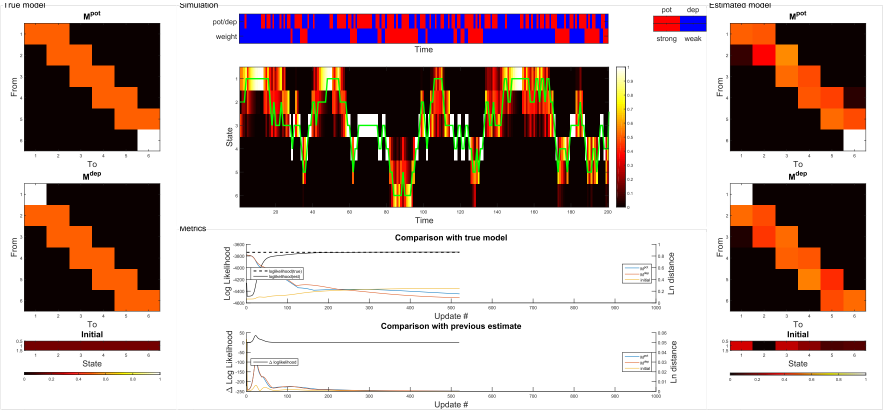


True model
Simulation
Estimated model
Metrics
pot/dep
weight
Time
pot
dep
strong
weak
State
From
To
Initial
Comparison with true model
Comparison with previous estimate
Log Likelihood
Δ Log Likelihood
Ln distance
Update #
loglikelihood(true)
loglikelihood(est)
Δ loglikelihood
initial

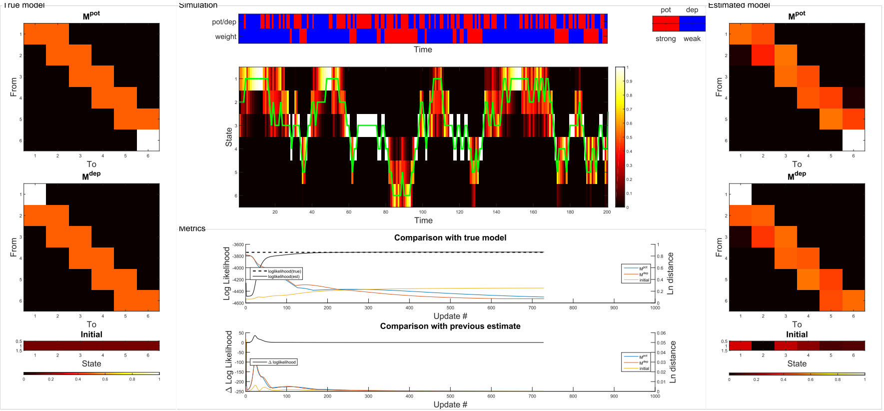
Fitting algorithm



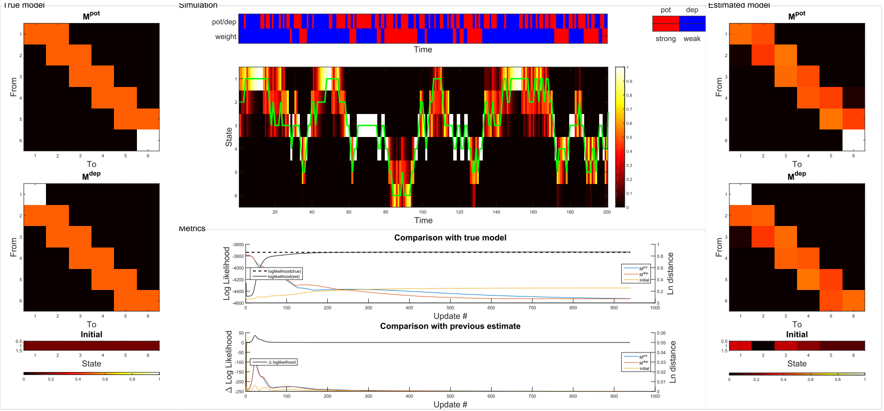
Fitting algorithm



Fitting algorithm



Fitting algorithm



Summary

- We have formulated a general theory of learning and memory with complex synapses.
- We find a memory envelope: a single curve that cannot be exceeded by the memory curve of *any* synaptic model.
- Synaptic complexity (M internal states) raises the memory envelope linearly in M for times $> \mathcal{O}(M^2)$.
- We understood which types of synaptic structure are useful for storing memories for different timescales.

Acknowledgements

Thanks to:

- Surya Ganguli
- Stefano Fusi
- Marcus Benna
- David Sussillo
- Jascha Sohl-Dickstein

References I

Larry R Squire and Pablo Alvarez.

"Retrograde amnesia and memory consolidation: a neurobiological perspective".

Current Opinion in Neurobiology, 5(2):169–177, (April, 1995) .

3

James L McClelland, Bruce L McNaughton, and Randall C O'Reilly.

"Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory.", 1995.

3

Jan Born and Ines Wilhelm.

"System consolidation of memory during sleep.".

Psychological research, 76(2):192–203, (mar, 2012) .

3

References II

Phillip J.E. Attwell, Samuel F. Cooke, and Christopher H. Yeo.
"Cerebellar Function in Consolidation of a Motor Memory".
Neuron, 34(6):1011–1020, (jun, 2002) .

3

Samuel F Cooke, Phillip J E Attwell, and Christopher H Yeo.
"Temporal properties of cerebellar-dependent memory consolidation.".
The Journal of neuroscience : the official journal of the Society for Neuroscience, 24(12):2934–41, (mar, 2004) .

3

Richard D. Emes and Seth G.N. Grant.
"Evolution of Synapse Complexity and Diversity".
Annual Review of Neuroscience, 35(1):111–131, (2012) .

3

References III

Carl C. H. Petersen, Robert C. Malenka, Roger A. Nicoll, and John J. Hopfield.

"All-or-none potentiation at CA3-CA1 synapses".

Proc. Natl. Acad. Sci. U.S.A., 95(8):4732–4737, (1998) .

4

Daniel H. O'Connor, Gayle M. Wittenberg, and Samuel S.-H. Wang.

"Graded bidirectional synaptic plasticity is composed of switch-like unitary events".

Proc. Natl. Acad. Sci. U.S.A., 102(27):9679–9684, (2005) .

4

D. J. Amit and S. Fusi.

"Constraints on learning in dynamic synapses".

Network: Computation in Neural Systems, 3(4):443–464, (1992) .

4

References IV

D. J. Amit and S. Fusi.
"Learning in neural networks with material synapses".
Neural Computation, 6(5):957–982, (1994) .
4

Friedrich T Sommer and Peter Dayan.
"Bayesian retrieval in associative memories with storage errors.".
IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council, 9(4):705–13, (jan, 1998) .
5 6 7 8 9 10 11 12 13

S. Fusi, P. J. Drew, and L. F. Abbott.
"Cascade models of synaptically stored memories".
Neuron, 45(4):599–611, (Feb, 2005) .
15

References V

S. Fusi and L. F. Abbott.

"Limits on the memory storage capacity of bounded synapses".

Nat. Neurosci., 10(4):485–493, (Apr, 2007) .

15

A. B. Barrett and M. C. van Rossum.

"Optimal learning rules for discrete synapses".

PLoS Comput. Biol., 4(11):e1000230, (Nov, 2008) .

15

LE Baum, T Petrie, George Soules, and Norman Weiss.

"A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains".

The annals of mathematical statistics, 41(1):164–171, (1970) .

42 43 44 45 46 47 48

References VI

Lawrence R Rabiner and Biing-Hwang Juang.

Fundamentals of speech recognition, volume 14 of *Signal Processing*.

Prentice Hall, Inc., Upper Saddle River, NJ, USA, 1993.

ISBN 0-13-015157-2.

42 43 44 45 46 47 48

A. P. Dempster, N. M. Laird, and D. B. Rubin.

"Maximum Likelihood from Incomplete Data via the EM Algorithm".

Journal of the Royal Statistical Society. Series B (Methodological), (October, 2007) .

42 43 44 45 46 47 48

J.G. Kemeny and J.L. Snell.

Finite markov chains.

Springer, 1960.

63

Technical detail: ordering states

Let $\mathbf{T}_{ij}$ = mean first passage time from state i to state j. Then:

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

is independent of the initial state i (Kemeney's constant).

[Kemeny and Snell (1960)]

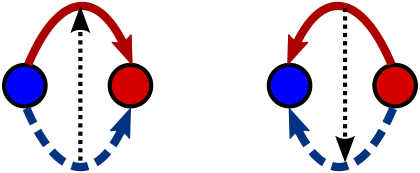
We define:

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

They can be used to arrange the states in an order (increasing η^- or decreasing η^+). back

Technical detail: upper/lower triangular

With states in order:



Endpoint: potentiation goes right, depression goes left.

back