# The synaptic dynamics of learning and memory

Subhaneil Lahiri

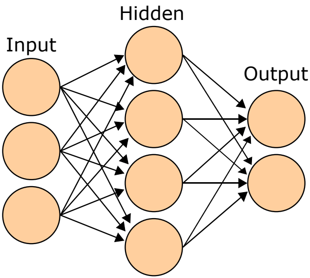
Stanford University, Applied Physics

May 13, 2019

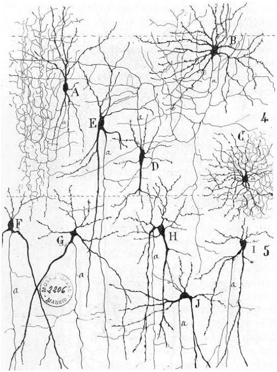
"A memory frontier for complex synapses", S Lahiri and S Ganguli.
*Adv. Neural Inf. Process. Syst. 26, pp. 1034–1042., (2013)*.

# What is a synapse?

Comp-neuro/deep-learning




Cellular biology



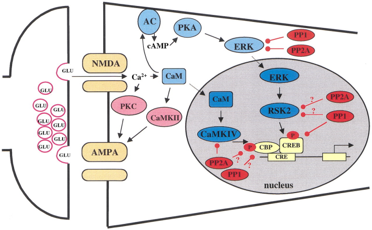

[Cajal (1899)]

# What is a synapse?

Comp-neuro/deep-learning

$$W_{ij}$$

Cellular biology



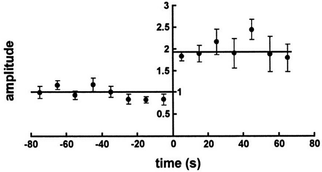

[Klann (2002)]

# Storage capacity of synaptic memory

Hopfield model, perceptron: capacity $\propto N$, (# synapses).

Assumes unbounded analogue synapses

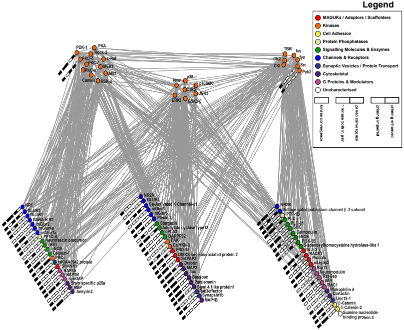
With discrete, finite synapses:
$\Longrightarrow$ memory capacity $\sim \mathcal{O}(\log N)$.
[Amit and Fusi (1992), Amit and Fusi (1994)]



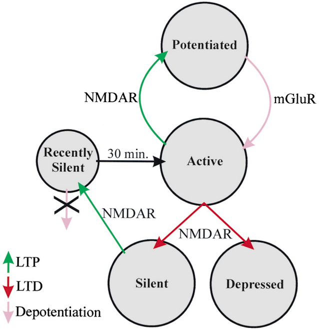

[Petersen et al. (1998), O'Connor et al. (2005)]

New memories overwrite old $\Longrightarrow$ stability-plasticity dilemma.

# Synapses are complex

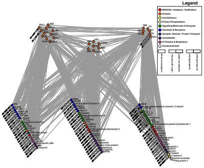



[Coba et al. (2009)]

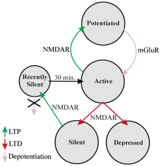



[Montgomery and Madison (2002)]

# Synapses are complex

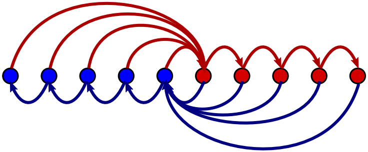


[Coba et al. (2009)]




[Montgomery and Madison (2002)]

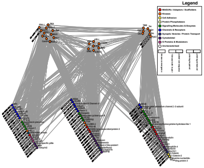
Cascade model



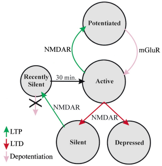
Capacity $\propto N^{2/3}$.

[Fusi et al. (2005)]

# Synapses are complex

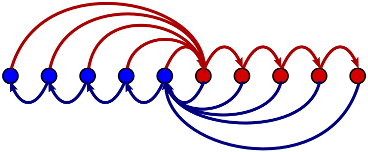



[Coba et al. (2009)]




[Montgomery and Madison (2002)]

Cascade model



Capacity $\propto N^{2/3}$. [Fusi et al. (2005)]

Capacity $\propto N$. [Benna and Fusi (2016)]

# My approach

We want to study the structure-function relationship of biological processes.

Not trying to build a *single* model.

Instead, we build a broad framework of models to find:

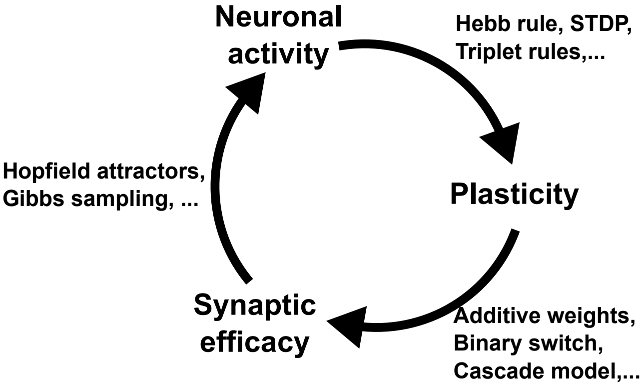
- $\rightarrow$ underlying mechanisms and principles.
- $\rightarrow$ trade-offs between aspects of performance (e.g. learning vs. memory).
- $\rightarrow$ properties of models that best manage these trade-offs.

# Outline

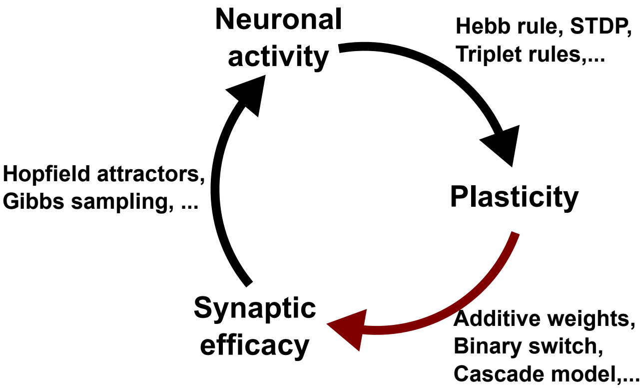
1. Timescales of memory
   - Quantifying memory quality
   - Frontiers of memory
   - Implications of memory limits

2. Designing experiments

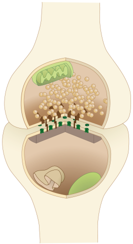
# Synaptic learning and memory



# Synaptic learning and memory

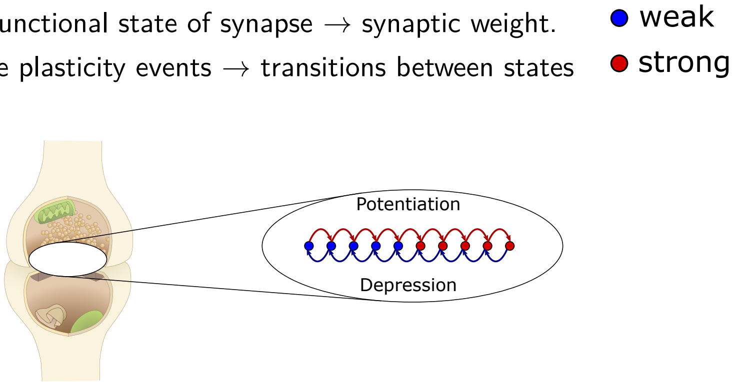


# Models of complex synaptic dynamics



# Models of complex synaptic dynamics

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



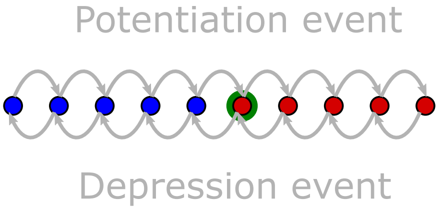

[Fusi et al. (2005), Fusi and Abbott (2007), Barrett and van Rossum (2008)]
[Smith et al. (2006), Lahiri and Ganguli (2013)]

# Models of complex synaptic dynamics

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

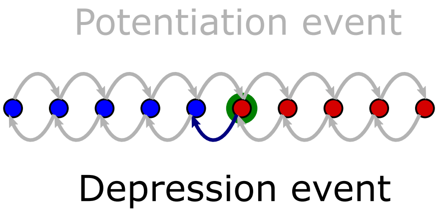
weak

strong



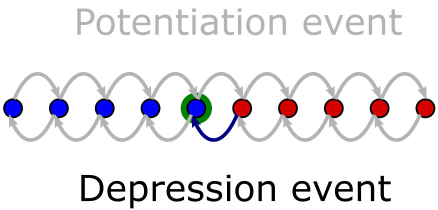
# Models of complex synaptic dynamics

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



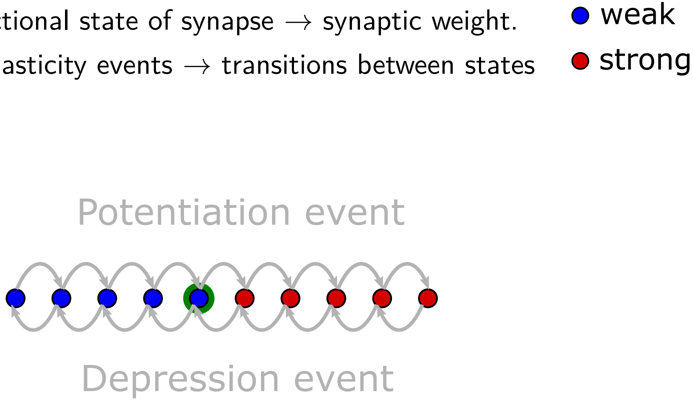
# Models of complex synaptic dynamics

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



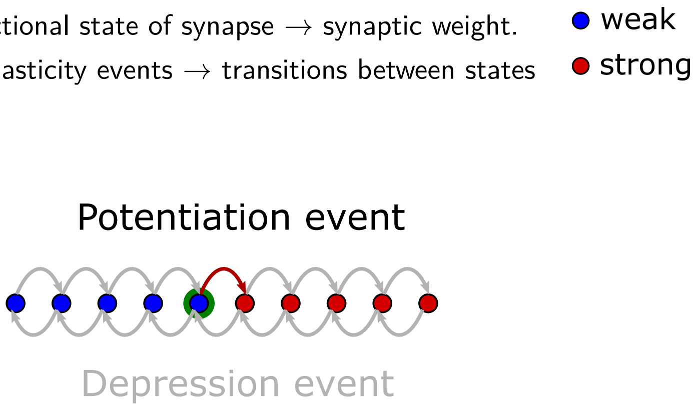
# Models of complex synaptic dynamics

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



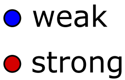
# Models of complex synaptic dynamics

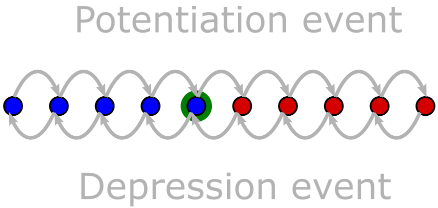
- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states



# Models of complex synaptic dynamics

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

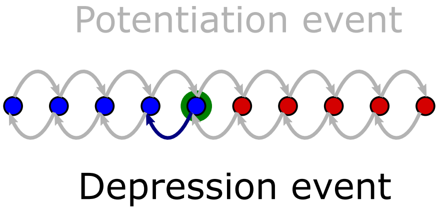


# Models of complex synaptic dynamics

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

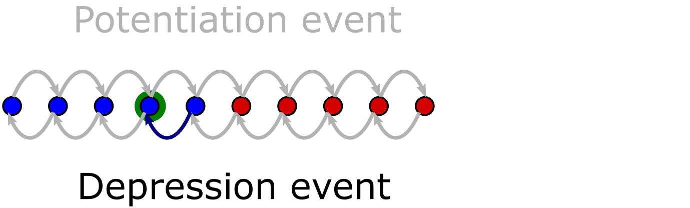
weak
strong




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

# Models of complex synaptic dynamics

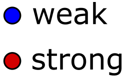
- Internal functional state of synapse → synaptic weight.
- Candidate plasticity events → transitions between states

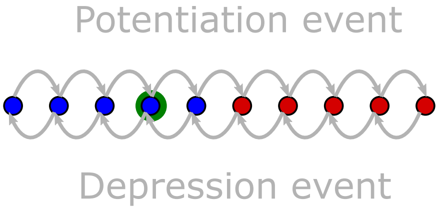



Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

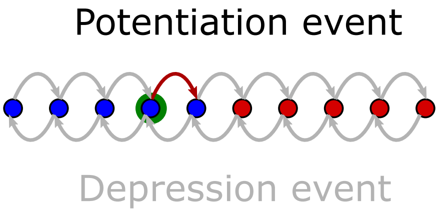



Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

# 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).

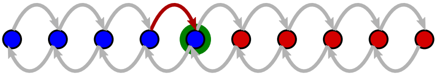
# Models of complex synaptic dynamics

- Internal functional state of synapse $\rightarrow$ synaptic weight.
- Candidate plasticity events $\rightarrow$ transitions between states

weak

strong

Potentiation event

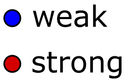


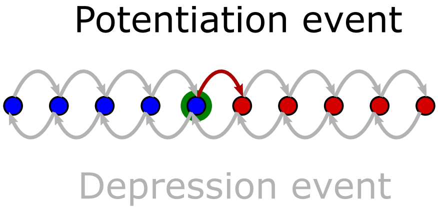
Depression event

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).

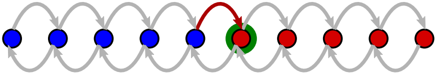
## Models of complex synaptic dynamics

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

weak

strong

Potentiation event



Depression event

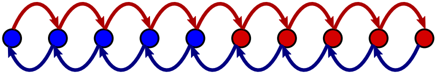
Metaplasticity: change propensity for plasticity
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak
strong

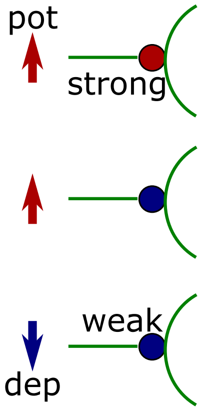
Potentiation



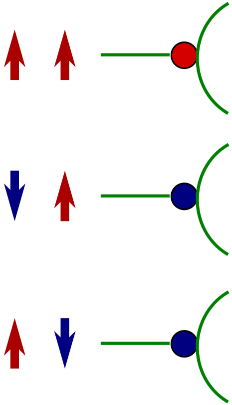
Depression

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

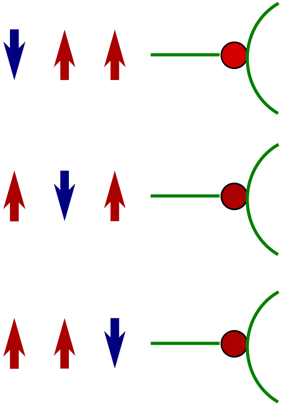
# Synaptic memory curves



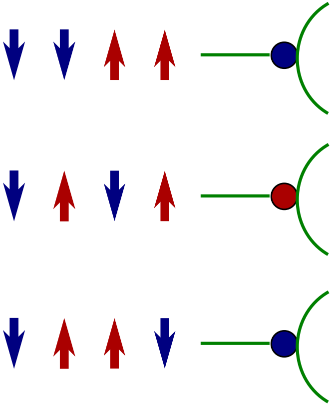
# Synaptic memory curves



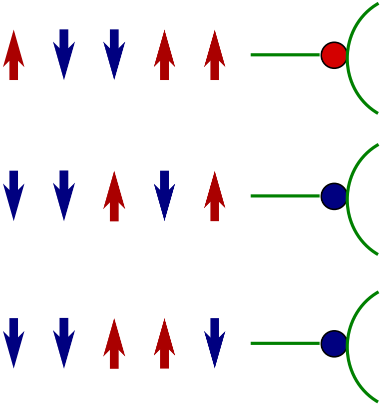
# Synaptic memory curves



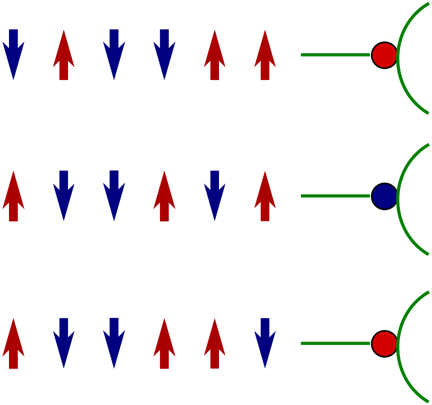
# Synaptic memory curves



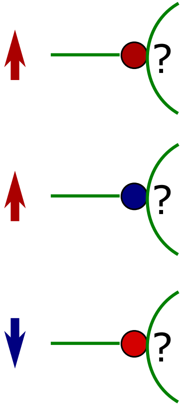
# Synaptic memory curves



# Synaptic memory curves



# Synaptic memory curves




Recognition memory: has this pattern been seen before?

Hypothesis test statistic: $\vec{w}(t) \cdot \vec{w}_{\text{test}}$. [Sommer and Dayan (1998)]

# Quantifying memory quality

Test if $\vec{w}(t) \cdot \vec{w}_{\text{test}} \gtrless$ threshold? [Sommer and Dayan (1998)]

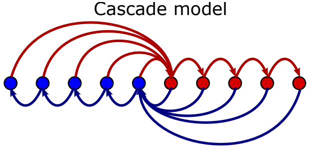
Compare with null-distribution: $P\left[\vec{w}_{\text{null}} \cdot \vec{w}_{\text{test}}\right]$

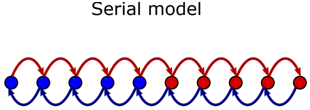
$$\text{SNR}(t) = \frac{\langle \vec{w}(t) \cdot \vec{w}_{\text{test}} \rangle - \langle \vec{w}_{\text{null}} \cdot \vec{w}_{\text{test}} \rangle}{\sqrt{\text{Var}(\vec{w}_{\text{null}} \cdot \vec{w}_{\text{test}})}},$$

$\Rightarrow$ discriminability: KL divergence, Chernoff distance, ROC curve, ...

# Specific models of complex synaptic dynamics

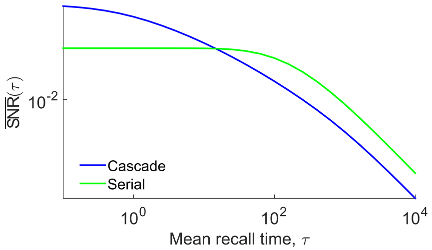
Two example models of complex synapses.



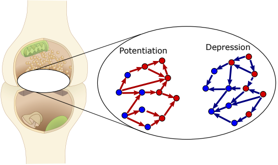

[Fusi et al. (2005), Leibold and Kempter (2008), Ben-Dayan Rubin and Fusi (2007)]

These have different memory storage properties

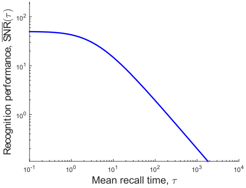


# General principles relating structure and function?

Synaptic structure




Synaptic function




- What are the fundamental limits of memory?
- Which models achieve these limits?
- What are the theoretical principles behind the optimal models?

# Parameters for synaptic dynamics

$$
\begin{aligned}
f^{\text{pot/dep}} &= \text{fraction of events that are pot/dep},\\
\text{pot. event: } \mathbf{M}^{\text{pot}}_{ij} &= \text{transition prob. } i \to j,\\
\text{dep. event: } \mathbf{M}^{\text{dep}}_{ij} &= \text{transition prob. } i \to j.
\end{aligned}
$$

Constraints:

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

# Eigenmodes are more convenient parameters

Eigenmodes:

$$\left(f^{\text{pot}}\mathbf{M}^{\text{pot}} + f^{\text{dep}}\mathbf{M}^{\text{dep}}\right)\mathbf{u}_a = \lambda_a \mathbf{u}_a.$$

Contribution to SNR:

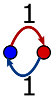
$$\text{SNR}(t) = \sum_a \mathcal{I}_a \, \mathrm{e}^{-t/\tau_a}.$$

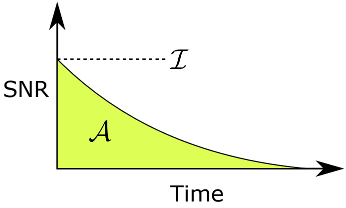
What are the constraints?

# Upper bounds on measures of memory

Initial SNR:

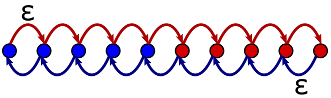
$$\mathcal{I} = \mathrm{SNR}(0) = \sum_a \mathcal{I}_a \le \sqrt{N}.$$




Area under curve:

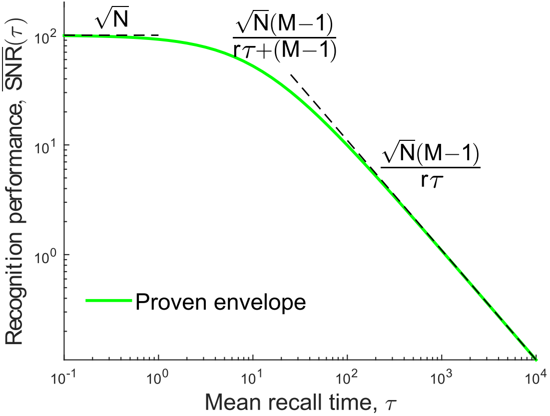
$$\mathcal{A} = \int_0^\infty \mathrm{SNR}(t)\,\mathrm{d}t = \sum_a \mathcal{I}_a \tau_a \le \sqrt{N}(M-1)/r.$$




[Lahiri and Ganguli (2013)]

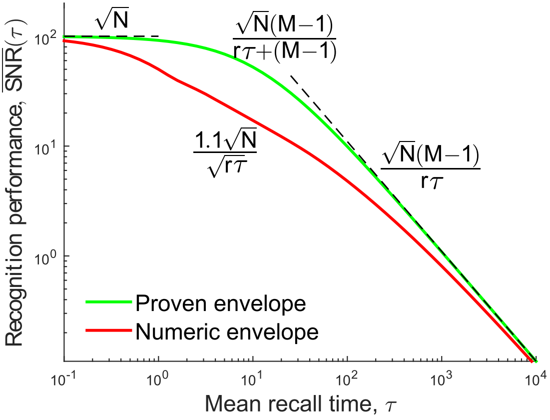
# Proven envelope: memory frontier

Upper bound on memory curve at *any* timescale.

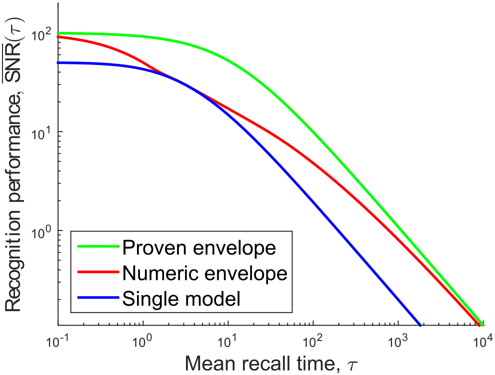


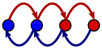
# Proven envelope: memory frontier

Upper bound on memory curve at *any* timescale.

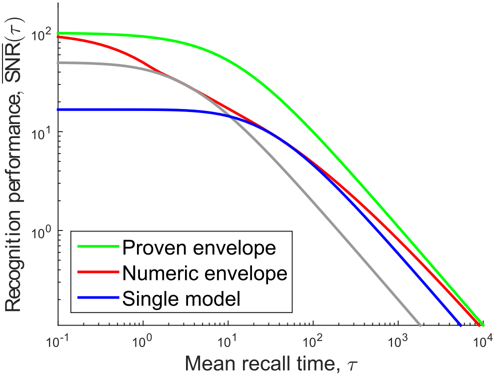


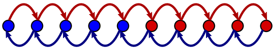
# Models that maximize memory for one timescale



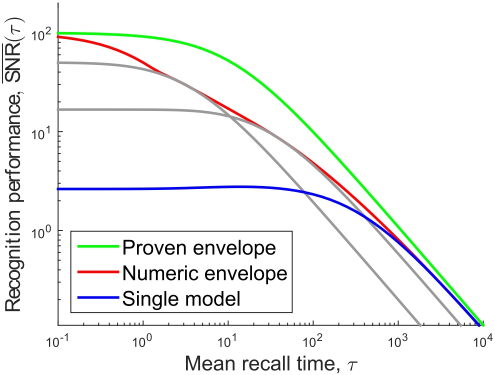



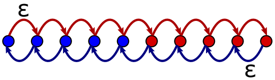
# Models that maximize memory for one timescale



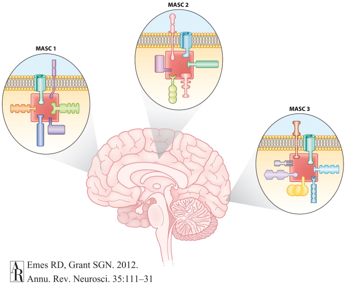
# Models that maximize memory for one timescale



# Synaptic diversity and timescales of memory

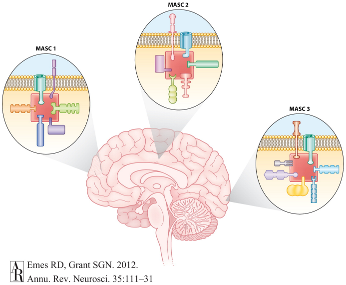
Different synapses have different molecular structures.




[Emes and Grant (2012)]

# Synaptic diversity and timescales of memory

Different synapses have different molecular structures.

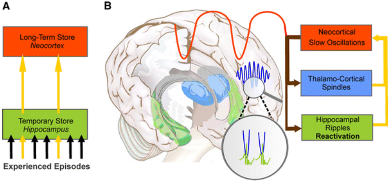



[Emes and Grant (2012)]

Memories stored in different places for different timescales

[Squire and Alvarez (1995)]

[McClelland et al. (1995)]




[Born and Wilhelm (2012)]

Also: Cerebellar cortex $\rightarrow$ nuclei.

[Attwell et al. (2002)]

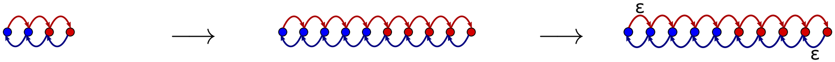
[Cooke et al. (2004)]

# Synaptic structure and function: general principles

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

What can we conclude?

Short timescale $\longrightarrow$ Moderate timescale $\longrightarrow$ Long timescale

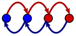
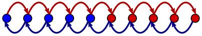
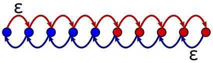


# Synaptic structure and function: general principles

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

What can we conclude?

Short timescale $\longrightarrow$ Moderate timescale $\longrightarrow$ Long timescale

 $\longrightarrow$  $\longrightarrow$ 

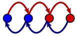
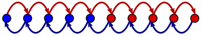
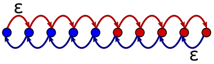

short topology $\longrightarrow$ long topology

# Synaptic structure and function: general principles

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

What can we conclude?

Short timescale $\longrightarrow$ Moderate timescale $\longrightarrow$ Long timescale

 $\longrightarrow$  $\longrightarrow$ 


short topology $\longrightarrow$ long topology

deterministic synapse $\longrightarrow$ stochastic synapse

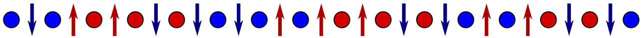
# Experimental tests?

Traditional experiments:

# Experimental tests?

Traditional experiments:

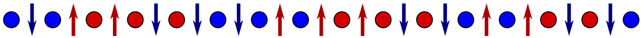
To fit a model: long sequence of small plasticity events.
Observe the changes in synaptic efficacy.



# Experimental tests?

Traditional experiments:

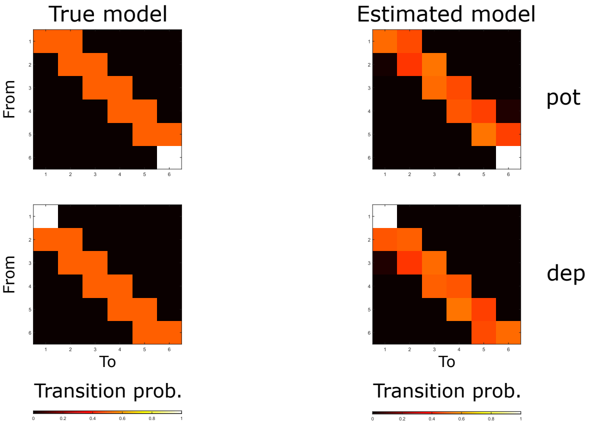
To fit a model: long sequence of small plasticity events.
Observe the changes in synaptic efficacy.



**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)]

# Simulated experiment




Problem: need *long* sequences.
Whole cell patch of postsynaptic neuron → Ca washout.

# Conclusions

- 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.
- We understood which types of synaptic structure are useful for storing memories for different timescales.
- We studied more than a single model. We studied *all possible models*, to extract general principles relating synaptic structure to function

# Acknowledgements

**Surya Ganguli**
Jascha Sohl-Dickstein
Friedemann Zenke
Sam Ocko
Stephane Deny
Jonathan Kadmon
Madhu Advani
Peiran Gao

David Sussillo

Niru Maheswaranathan
Ben Poole
Kiah Hardcastle
Lane McIntosh
Alex Williams
Christopher Stock
Sarah Harvey
Aran Nayebi

Stefano Fusi

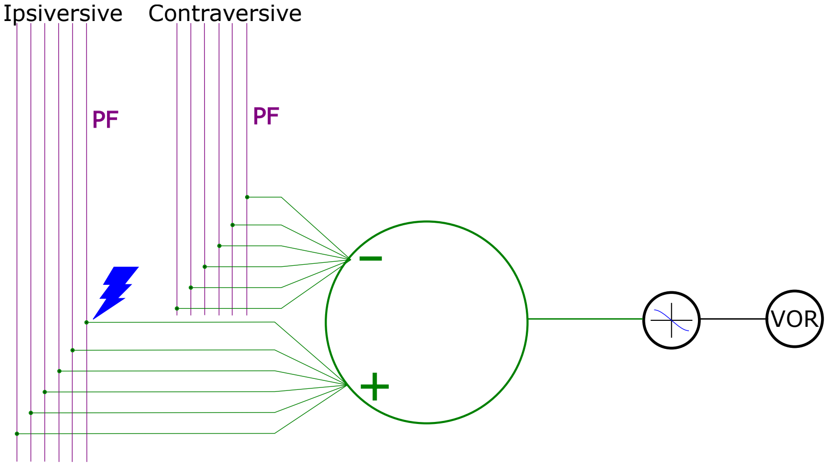
**Jennifer Raymond**
Barbara Nguyen-Vu
Grace Zhao
Aparna Suvrathan
Rhea Kimpo

**Carla Shatz**
Hanmi Lee

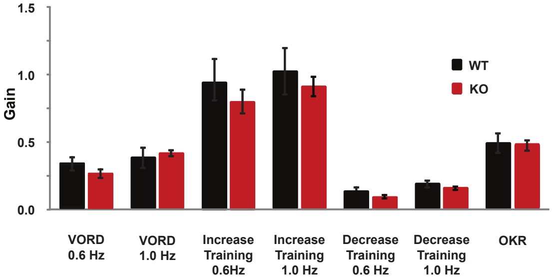
Marcus Benna

**Funding**: Swartz Foundation, Stanford Bio-X Genentech fellowship.

# Model of circuit

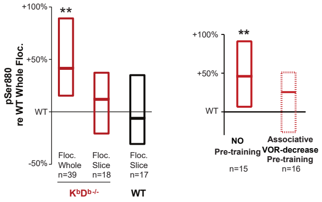


# Baseline



1.5
1.0
0.5
0.0
Gain
WT
KO
VORD 0.6 Hz
VORD 1.0 Hz
Increase Training 0.6Hz
Increase Training 1.0 Hz
Decrease Training 0.6 Hz
Decrease Training 1.0 Hz
OKR

# Evidence: level of depression



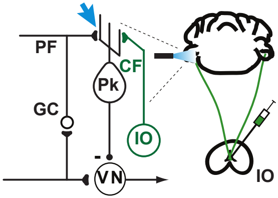

Basal level of GluR2 phosphorylation at serine 880 in AMPA receptor.

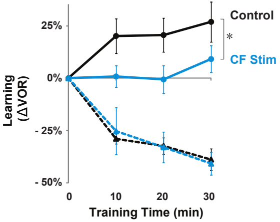
Biochemical signature of PF-Pk LTD.

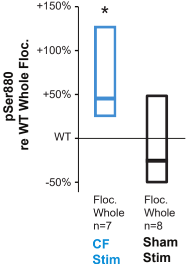
Shows that # depressed synapses in flocculus is larger in KO than WT.

# Evidence: saturation by CF stimulation

Use Channelrhodopsin to stimulate CF →increase LTD in PF-Pk synapses
→simulate saturation in WT.

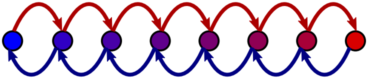


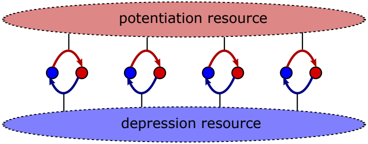



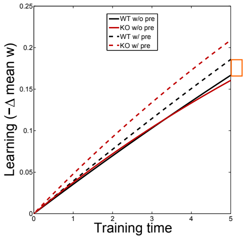
# Other models that fail

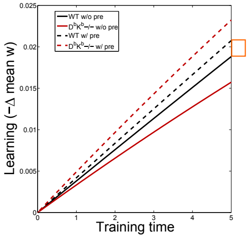
## Multistate synapse



Pooled resource model



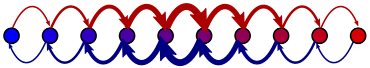



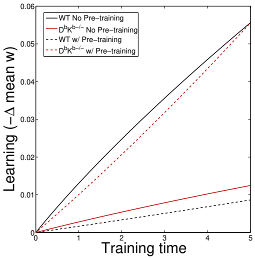




[Amit and Fusi (1994)]

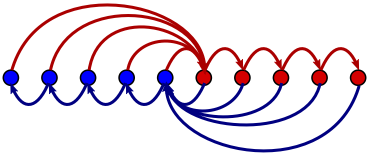
# Other models that work

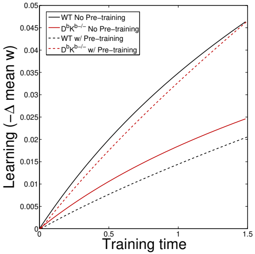
Non-uniform multistate model





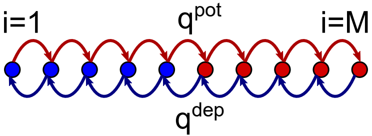

Cascade model






[Fusi et al. (2005)]

# Mathematical explanation




Serial synapse: $\boldsymbol{\pi}_i \sim \mathcal{N}\left(\frac{q^{\text{pot}}}{q^{\text{dep}}}\right)^i$.

Learning rate $\sim \boldsymbol{\pi}_{M/2}\left(\frac{q^{\text{dep}}}{q^{\text{pot}}}\right) = \mathcal{N}\left(\frac{q^{\text{pot}}}{q^{\text{dep}}}\right)^{\frac{M}{2}-1}$.

For $M > 2$: larger $q^{\text{dep}} \implies$ slower learning.

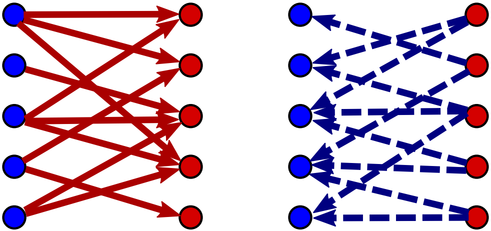
For $M = 2$: larger $q^{\text{dep}} \implies$ larger $\mathcal{N} \implies$ faster learning.

# Initial SNR as flux

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

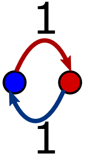
$$\mathrm{SNR}(0) \leq \frac{4\sqrt{N}}{r}\,\mathbf{\Phi}_{-+}.$$

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



# Two-state model

Two-state model equivalent to previous slide:

Transitions: 


$$\Longrightarrow \quad \mathrm{SNR}(t) = \sqrt{N}\,(4 f^{\mathrm{pot}} f^{\mathrm{dep}})\,\mathrm{e}^{-rt}.$$

Maximal initial SNR:

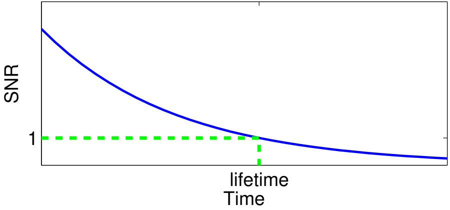
$$\mathrm{SNR}(0) \leq \sqrt{N}.$$

# Area under memory curve

$$\mathcal{A} = \int_0^\infty \mathrm{d}t \; \mathrm{SNR}(t), \qquad \overline{\mathrm{SNR}}(\tau) \to \frac{\mathcal{A}}{\tau} \quad \text{as} \quad \tau \to \infty.$$

Area bounds memory lifetime:

$$\begin{aligned} \mathrm{SNR}(\text{lifetime}) &= 1 \\ \Longrightarrow \qquad \text{lifetime} &< \mathcal{A}. \end{aligned}$$




This area has an upper bound:

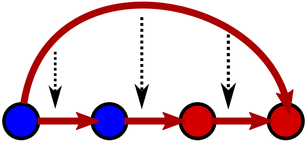
$$\mathcal{A} \leq \sqrt{N}(M-1)/r.$$

Saturated by a model with linear chain topology.

# Proof of area bound

For any model, we can construct perturbations that

- preserve equilibrium distribution,
- increase area.



details

e.g. decrease "shortcut" transitions, increase bypassed "direct" ones.
Endpoint: linear chain

The area of this model is

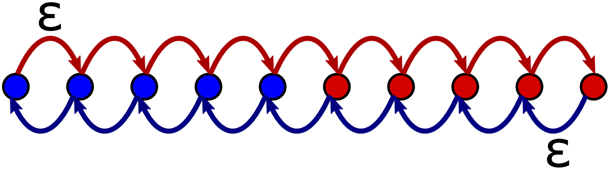
$$A = \frac{2\sqrt{N}}{r} \sum_k \pi_k \left| k - \langle k \rangle \right|.$$

Max: equilibrium probability distribution concentrated at both ends.

[Barrett and van Rossum (2008)]

# Saturating model

Make end states “sticky”




Has long decay time, but terrible initial SNR.

$$\lim_{\varepsilon \to 0} A = \sqrt{N}(M-1)/r.$$

# Technical detail: ordering states

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

$$\eta = \sum_j \mathbf{T}_{ij} \boldsymbol{\pi}_j,$$

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

[Kemeny and Snell (1960)]

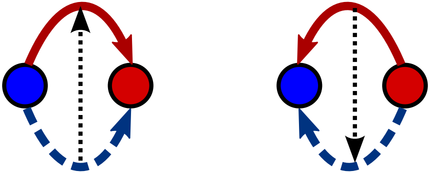
We define:

$$\eta_i^+ = \sum_{j \in \text{strong}} \mathbf{T}_{ij} \boldsymbol{\pi}_j, \qquad \eta_i^- = \sum_{j \in \text{weak}} \mathbf{T}_{ij} \boldsymbol{\pi}_j.$$

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

# Technical detail: upper/lower triangular

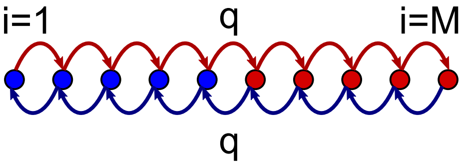
With states in order:



Endpoint: potentiation goes right, depression goes left.

back

# Intuition for using topology




$$\mathcal{I} \propto q, \qquad \max_a \tau_a \propto \frac{1}{q},$$

$$\mathcal{I} \propto \frac{1}{M}, \qquad \max_a \tau_a \propto M^2,$$

$$\Longrightarrow$$

Stochasticity: $\mathcal{I} \propto \dfrac{1}{\tau_{\max}}$,

Topology: $\mathcal{I} \propto \dfrac{1}{\sqrt{\tau_{\max}}}$.

# References I


Eric Klann.
"Metaplastic Protein Phosphatases".
*Learning and Memory*, 9(4):153–155, (2002) , http://learnmem.cshlp.org/content/9/4/153.full.pdf+html.
3

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

D. J. Amit and S. Fusi.
"Learning in neural networks with material synapses".
*Neural Comput.*, 6(5):957–982, (1994) .
4 61

# References II


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

# References III


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) .

5 6 7

Johanna M. Montgomery and Daniel V. Madison.

"State-Dependent Heterogeneity in Synaptic Depression between Pyramidal Cell Pairs".

*Neuron*, 33(5):765 – 777, (2002) .

5 6 7

# References IV


S. Fusi, P. J. Drew, and L. F. Abbott.
"Cascade models of synaptically stored memories".
*Neuron*, 45(4):599–611, (February, 2005) .
6 7 13 36 62

Marcus K. Benna and Stefano Fusi.
"Computational principles of synaptic memory consolidation".
*Nature Neuroscience*, 19(12):1697–1706, (July, 2016) , arXiv:1507.07580 [q-bio.NC].
6 7

S. Fusi and L. F. Abbott.
"Limits on the memory storage capacity of bounded synapses".
*Nat. Neurosci.*, 10(4):485–493, (Apr, 2007) .
13

# References V


A. B. Barrett and M. C. van Rossum.
"Optimal learning rules for discrete synapses".
*PLoS Comput. Biol.*, 4(11):e1000230, (November, 2008) .
13 67

Maurice A Smith, Ali Ghazizadeh, and Reza Shadmehr.
"Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning".
*PLoS Biol.*, 4(6):e179, (May, 2006) .
13

# References VI


Subhaneil Lahiri and Surya Ganguli.

"A memory frontier for complex synapses".

In C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Adv. Neural Inf. Process. Syst. 26*, pages 1034–1042. NIPS, 2013.

URL `https://papers.nips.cc/paper/4872-a-memory-frontier-for-complex-synapses`.

13 40

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) .

34 35

# References VII


Christian Leibold and Richard Kempter.

"Sparseness Constrains the Prolongation of Memory Lifetime via Synaptic Metaplasticity".

*Cereb. Cortex*, 18(1):67–77, (2008) .

36

Daniel D Ben-Dayan Rubin and Stefano Fusi.

"Long memory lifetimes require complex synapses and limited sparseness".

*Front. Comput. Neurosci.*, 1:1–14, (November, 2007) .

36

Richard D. Emes and Seth G.N. Grant.

"Evolution of Synapse Complexity and Diversity".

*Annual Review of Neuroscience*, 35(1):111–131, (2012) .

46 47

# References VIII


Larry R Squire and Pablo Alvarez.
"Retrograde amnesia and memory consolidation: a neurobiological perspective".
*Current Opinion in Neurobiology*, 5(2):169–177, (April, 1995) .
46 47

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.
46 47

Jan Born and Ines Wilhelm.
"System consolidation of memory during sleep.".
*Psychological research*, 76(2):192–203, (mar, 2012) .
46 47

# References IX


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) .
46 47

Samuel F Cooke, Phillip J E Attwell, and Christopher H Yeo.
"Temporal properties of cerebellar-dependent memory consolidation.".
*J. Neurosci.*, 24(12):2934–41, (mar, 2004) .
46 47

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) .
51 52 53

# References X


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.
51 52 53

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) .
51 52 53

J.G. Kemeny and J.L. Snell.
*Finite markov chains*.
Springer, 1960.
69