

# Learning and memory with complex synaptic plasticity

Subhaneil Lahiri

Stanford University, Applied Physics

May 8, 2019

# My research areas

## **Learning and memory**

Structure of synapses & function.  
Learning v. remembering tradeoff.  
Success & failure in trying to enhance learning.

## **Energy use in living systems**

Energy cost of signalling/sensing.  
Tradeoffs with accuracy & speed.  
Thermodynamics  $\leftrightarrow$  information geometry.

## **High dimensional statistics**

Theory of random projections and the geometry of data.  
Neural recordings as projections.

# What is a synapse?

Theorists

Experimentalists

# What is a synapse?

Theorists

Experimentalists

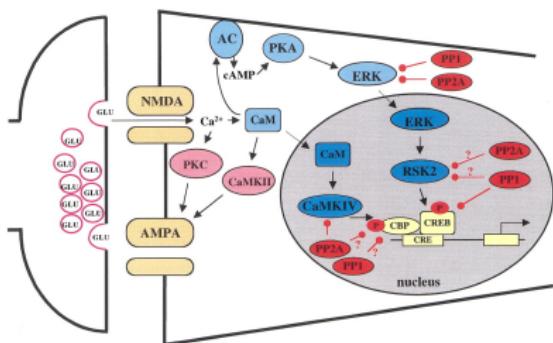
$$W_{ij}$$

# What is a synapse?

Theorists

$W_{ij}$

Experimentalists



[Klann (2002)]

# Storage capacity of synaptic memory

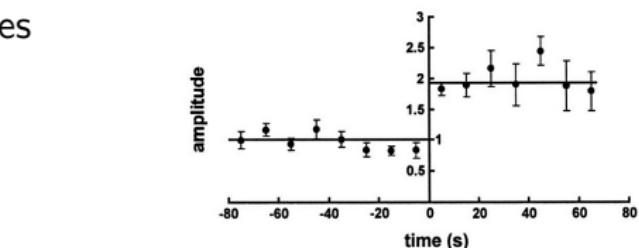
Hopfield, perceptron have capacity  $\propto N$ , (# synapses).

Assumes unbounded analogue synapses

With discrete, finite synapses:

$\implies$  memory capacity  $\sim \mathcal{O}(\log N)$ .

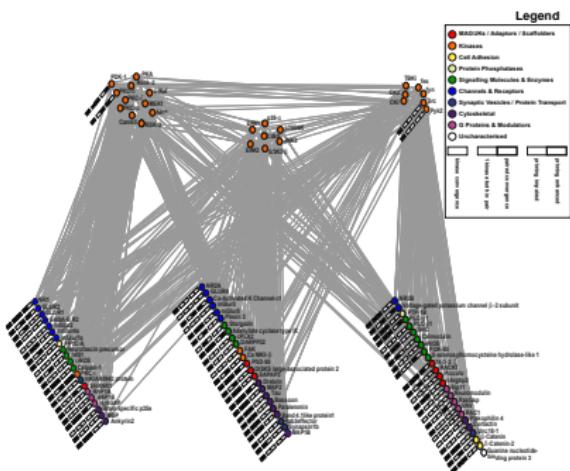
[Amit and Fusi (1992), Amit and Fusi (1994)]



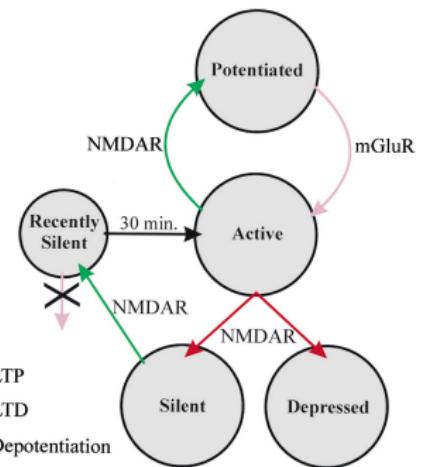
[Petersen et al. (1998), O'Connor et al. (2005)]

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

## Synapses are complex

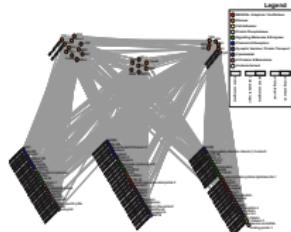


[Coba et al. (2009)]

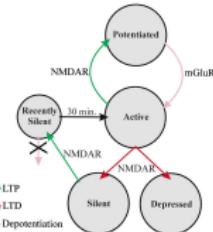


[Montgomery and Madison (2002)]

# Synapses are complex

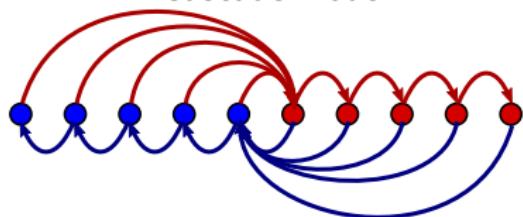


[Coba et al. (2009)]



[Montgomery and Madison (2002)]

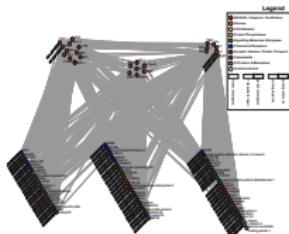
Cascade model



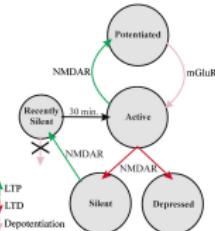
Capacity  $\propto N^{2/3}$ .

[Fusi et al. (2005)]

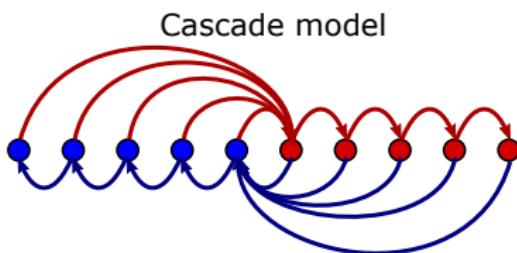
# Synapses are complex



[Coba et al. (2009)]



[Montgomery and Madison (2002)]



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:

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

# Outline

## 1 Memory over different timescales

- Quantifying memory quality
- Frontiers of memory
- Implications of memory limits

## 2 Learning with enhanced plasticity

- Effects of enhanced plasticity on cerebellar learning
- Synaptic models of cerebellar learning
- Learning outcomes of mice and models

## 3 Designing experiments

# Section 1

## Memory over different timescales

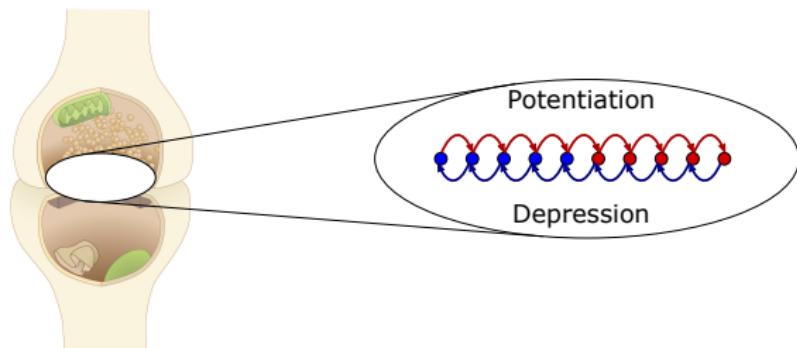
“A memory frontier for complex synapses”, S Lahiri and S Ganguli.  
*Adv. Neural Inf. Process. Syst. 26, pp. 1034–1042., (2013).*  
NeurIPS 2013 Outstanding Paper Award.

# 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,...

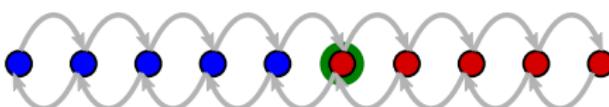
[Fusi et al. (2005), Fusi and Abbott (2007), Barrett and van Rossum (2008)]

[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

Potentiation event

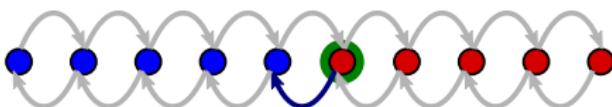


Depression event

# Models of complex synaptic dynamics

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

Potentiation event



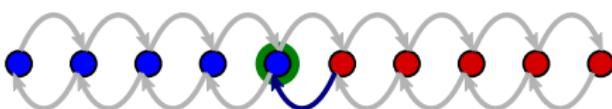
Depression event

# Models of complex synaptic dynamics

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

weak  
strong

Potentiation event

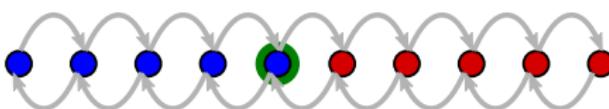


Depression event

# Models of complex synaptic dynamics

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

Potentiation event

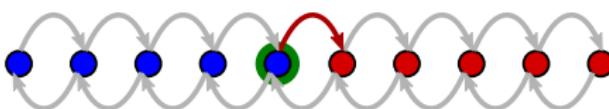


Depression event

# Models of complex synaptic dynamics

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

Potentiation event

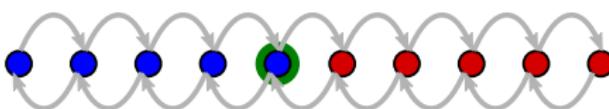


Depression event

# Models of complex synaptic dynamics

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

Potentiation event



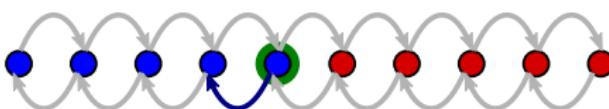
Depression event

# Models of complex synaptic dynamics

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

weak  
strong

Potentiation event



Depression event

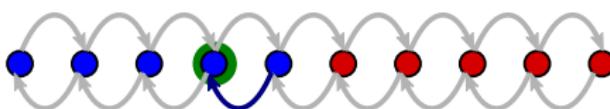
Metaplasticity: change propensity for plasticity  
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak  
strong

Potentiation event



Depression event

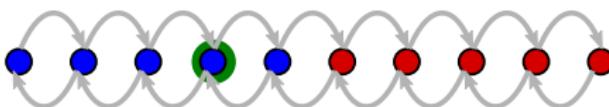
Metaplasticity: change propensity for plasticity  
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak  
strong

Potentiation event



Depression event

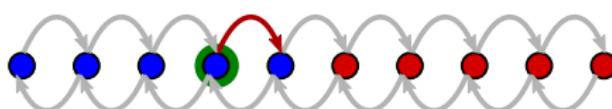
Metaplasticity: change propensity for plasticity  
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak  
strong

## Potentiation event



## Depression event

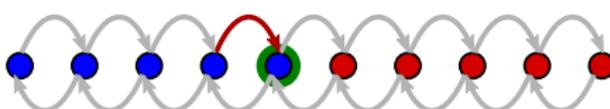
Metaplasticity: change propensity for plasticity  
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak  
strong

## Potentiation event



## Depression event

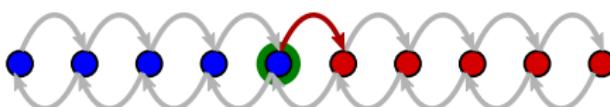
Metaplasticity: change propensity for plasticity  
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak  
strong

## Potentiation event



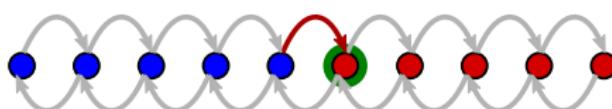
## Depression event

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 event



## Depression event

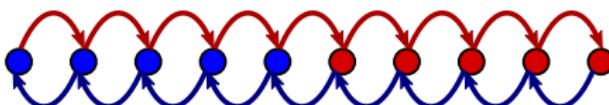
Metaplasticity: change propensity for plasticity  
(independent of change in synaptic weight).

# Models of complex synaptic dynamics

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

weak  
strong

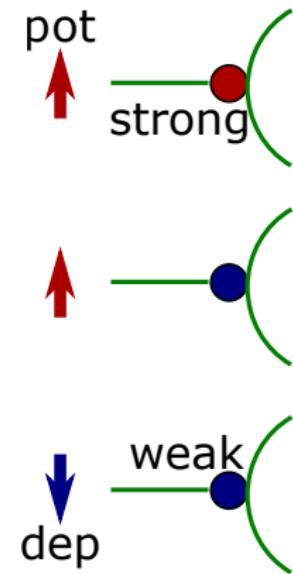
## Potentiation



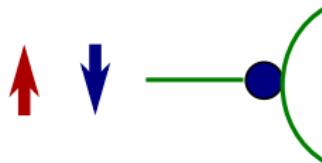
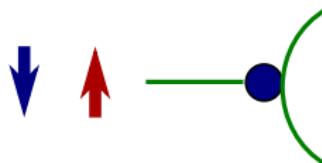
## Depression

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

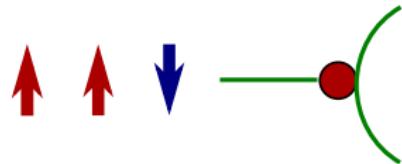
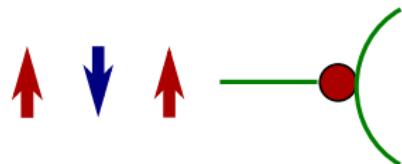
## Synaptic memory curves



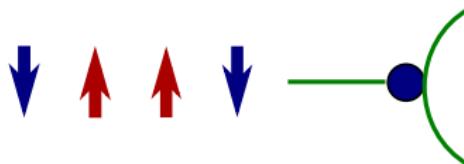
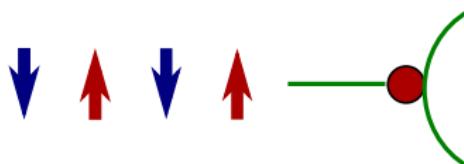
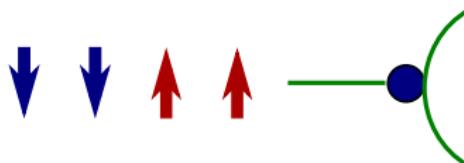
## Synaptic memory curves



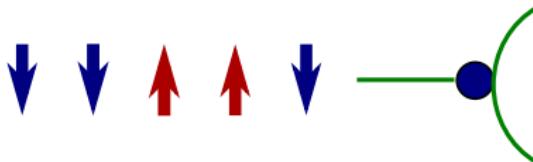
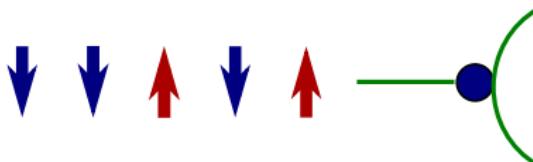
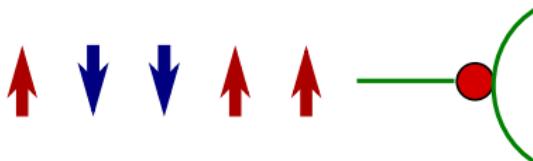
## Synaptic memory curves



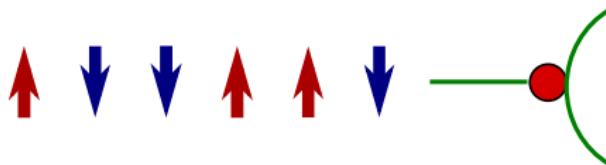
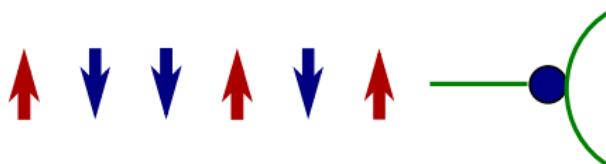
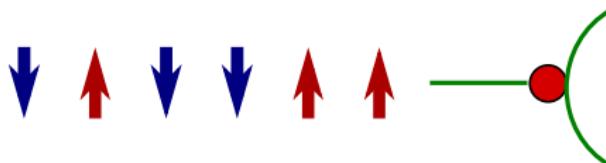
## Synaptic memory curves



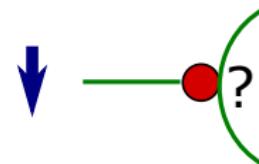
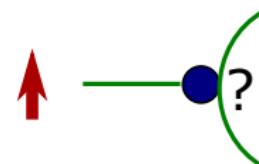
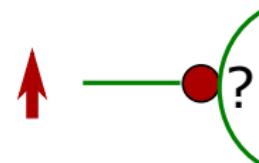
## Synaptic memory curves



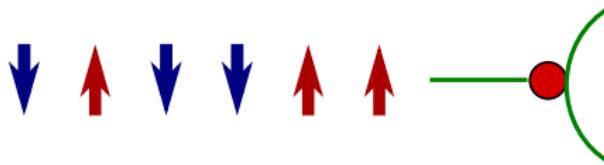
## Synaptic memory curves



## Synaptic memory curves



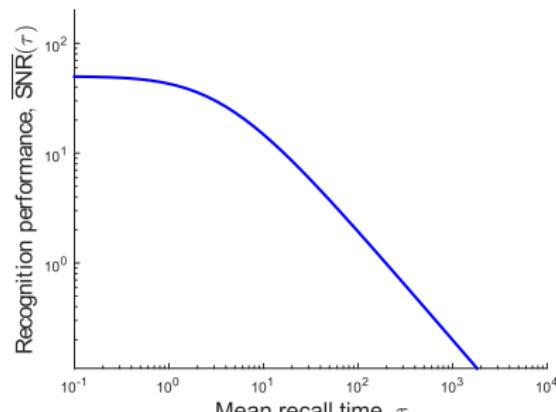
# Synaptic memory curves



Synapses store a sequence of memories.

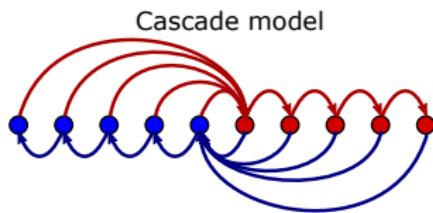
Recognition memory: has this pattern been seen before?

Performance described by SNR of  $\vec{w}(t) \cdot \vec{w}_{\text{test}}$ .

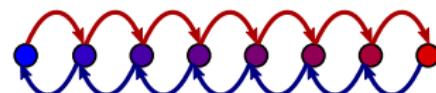


# Specific models of complex synaptic dynamics

## Two example models of complex synapses.

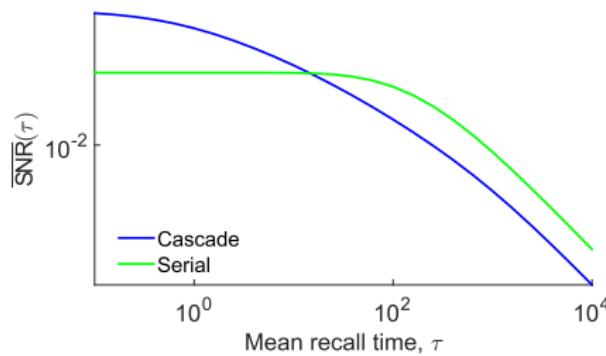


## Multistate synapse



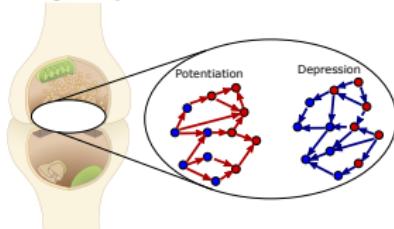
[Fusi et al. (2005), Leibold and Kemper (2008), Ben-Dayan Rubin and Fusi (2007)]

These have different memory storage properties

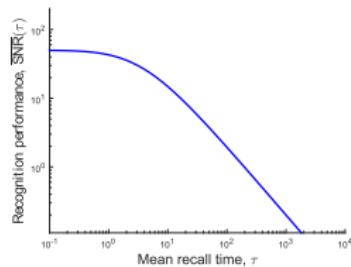


# General principles relating structure and function?

## Synaptic structure



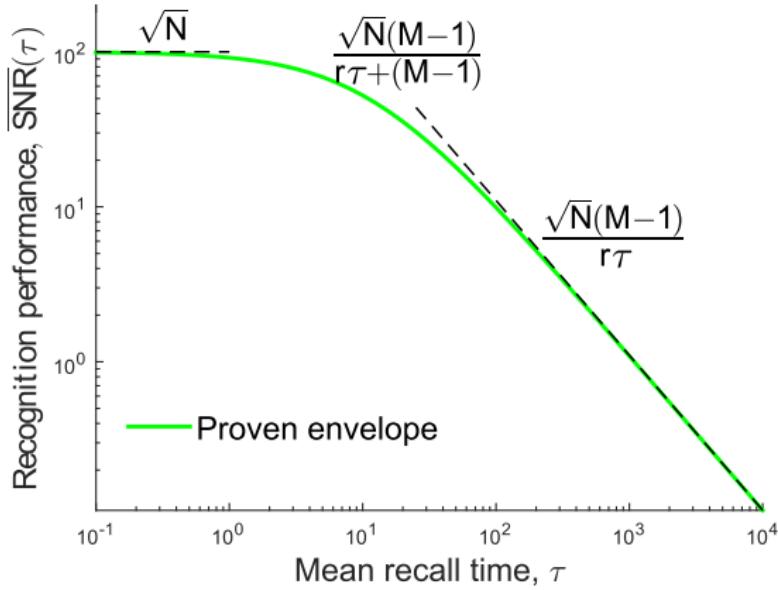
## Synaptic function



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

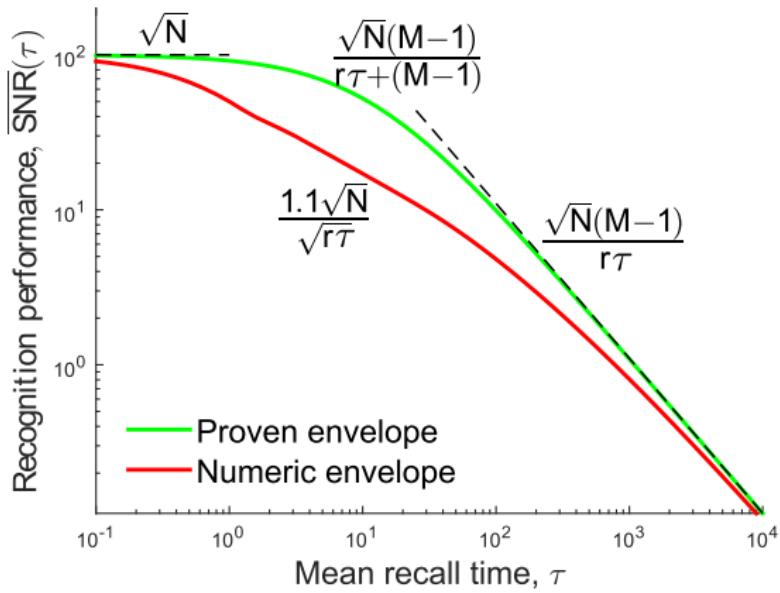
## Proven envelope: memory frontier

Upper bound on memory curve at *any* timescale.

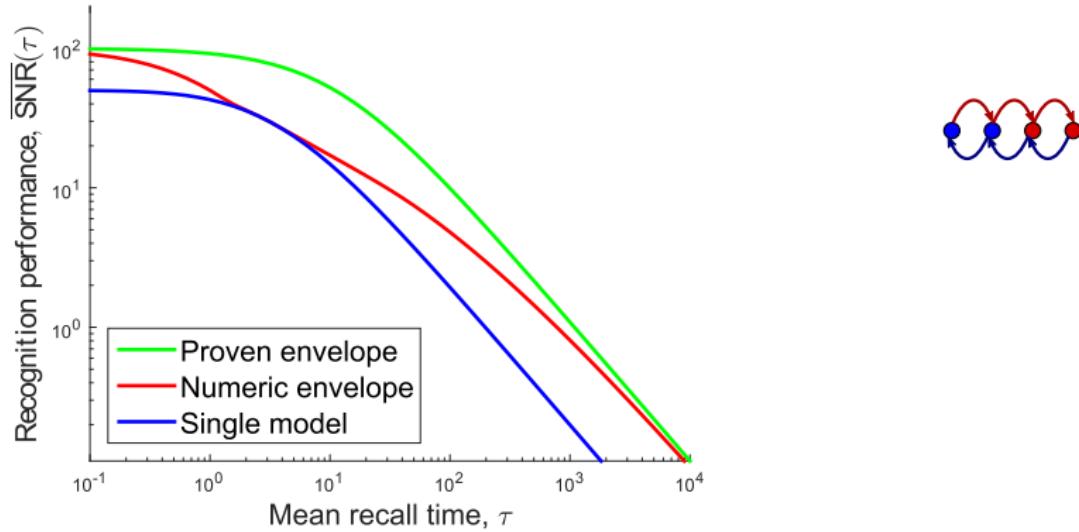


## Proven envelope: memory frontier

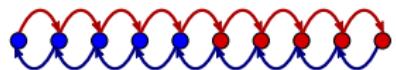
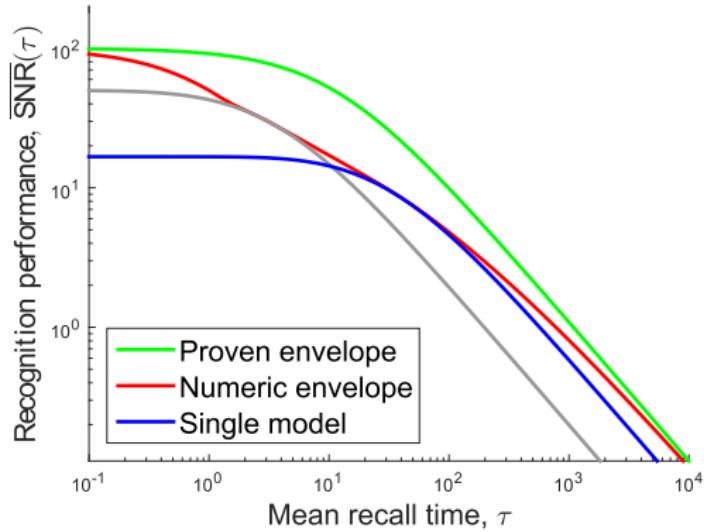
Upper bound on memory curve at *any* timescale.



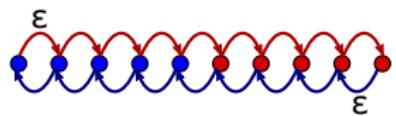
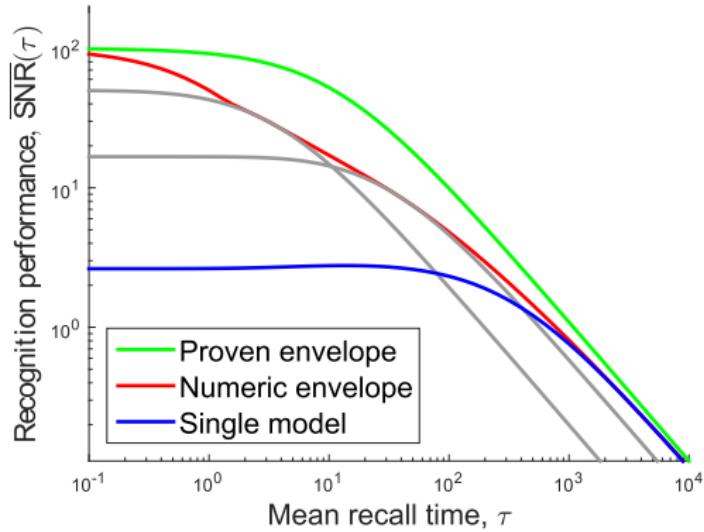
# Models that maximize memory for one timescale



## Models that maximize memory for one timescale

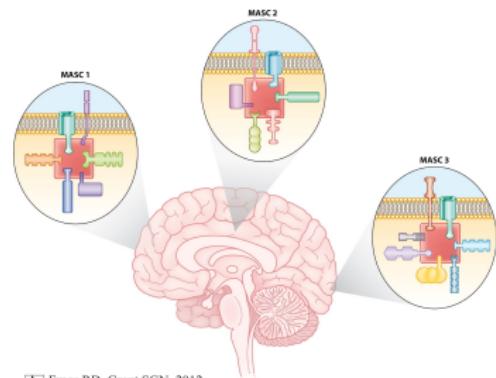


# Models that maximize memory for one timescale



# Synaptic diversity and timescales of memory

Different synapses have different molecular structures.

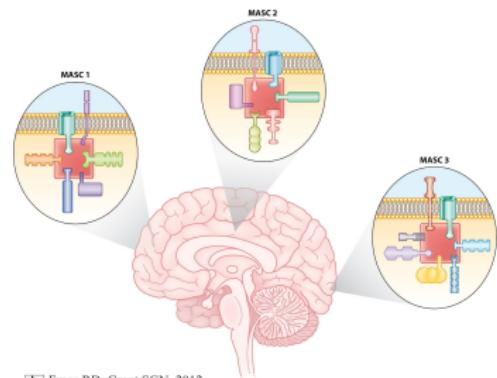


[A] Emes RD, Grant SGN. 2012.  
Annu. Rev. Neurosci. 35:111–31

[Emes and Grant (2012)]

# Synaptic diversity and timescales of memory

Different synapses have different molecular structures.



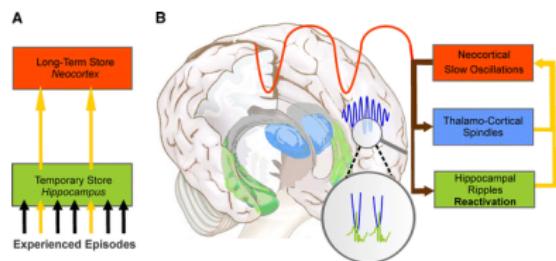
[Emes RD, Grant SGN. 2012.  
Annu. Rev. Neurosci. 35:111–31]

[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 → nuclei.

[Attwell et al. (2002)]

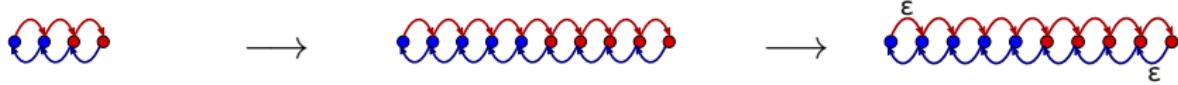
[Cooke et al. (2004)]

# Synaptic structure and function: general principles

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

What can we conclude?

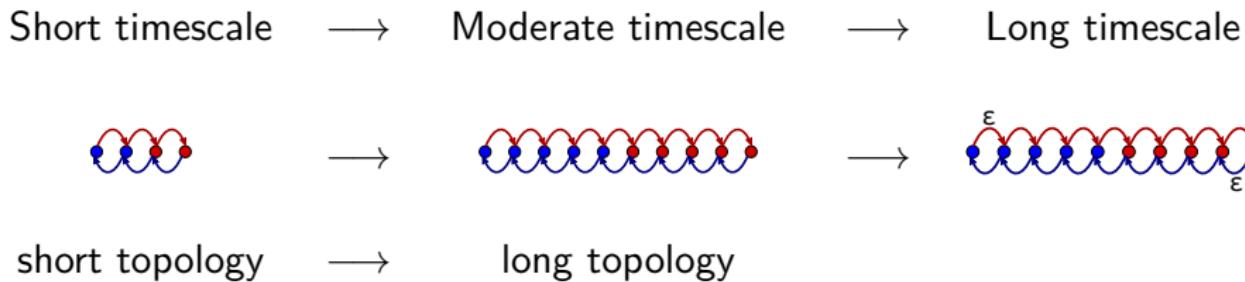
Short timescale → Moderate timescale → Long timescale



# Synaptic structure and function: general principles

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

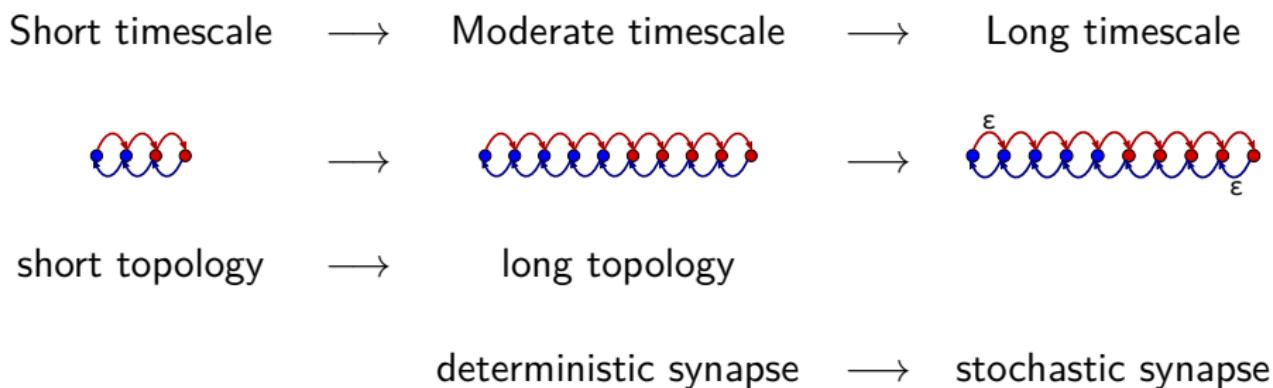
What can we conclude?



# Synaptic structure and function: general principles

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

What can we conclude?



# 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

## Section 2

### Learning with enhanced plasticity

"A saturation hypothesis to explain both enhanced and impaired learning with enhanced plasticity", TDB Nguyen-Vu, GQ Zhao, S Lahiri, RR Kimpo, H Lee, S Ganguli, CJ Shatz, JL Raymond.  
*eLife*, 6:e20147, (Feb., 2017).

Highest rated abstract at Cosyne 2014.

# Benefits of enhanced plasticity?

Learning requires synaptic plasticity.

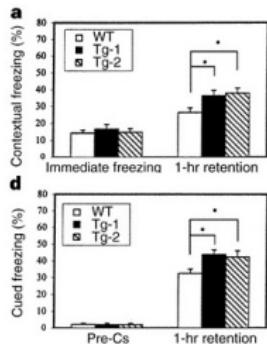


Can we enhance learning by enhancing plasticity?

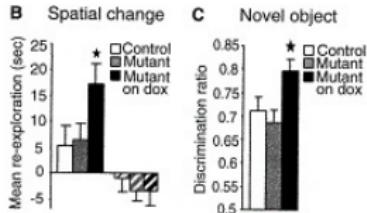


# Enhanced plasticity can enhance learning

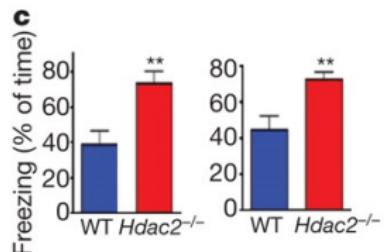
## Overexpress NR2B



## Inhibit CN



## Knockout Hdac2



## Fear conditioning

[Tang et al. (1999)]

## Novel object recog.

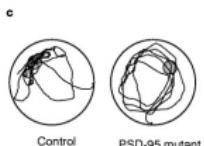
[Malleret et al. (2001)]

## Fear conditioning

[Guan et al. (2009)]

# Enhanced plasticity can *impair* learning

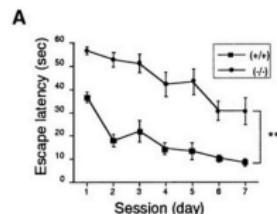
Mutate PSD-95



Water maze

[Migaud et al. (1998)]

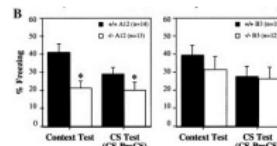
Knockout PTP $\delta$



Water maze

[Uetani et al. (2000)]

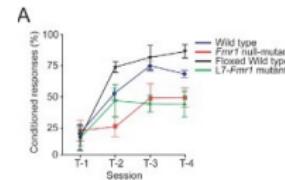
Delete Tmod2



Fear cond.

[Cox et al. (2003)]

Knockout FMR1



Eyeblink

[Koekkoek et al. (2005)]

also: [Hayashi et al. (2004), Rutten et al. (2008)]

# Overview

Sometimes enhanced plasticity → enhanced learning.  
Sometimes enhanced plasticity → impaired learning.

Why? How? When?

# Overview

Sometimes enhanced plasticity → enhanced learning.  
Sometimes enhanced plasticity → impaired learning.

Why? How? When?

Mice with enhanced cerebellar plasticity can show **both** impaired and enhanced learning.

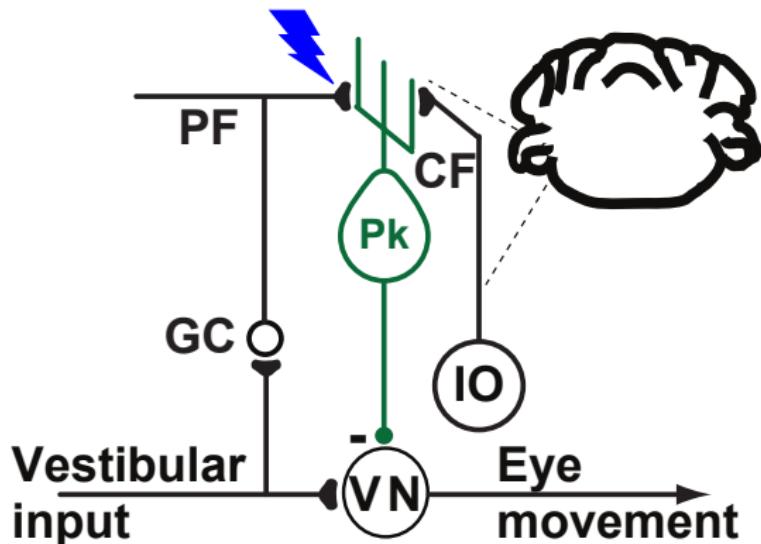
Simple synapses **cannot** explain behaviour. **Complex synapses** are required.  
→ predictions for synaptic physiology.

# Vestibulo-Occular Reflex training

VOR Increase  
Training



VOR Decrease  
Training



VOR increase: LTD in PF-Pk synapses.

VOR decrease: different mechanism,  
also reverses LTD in PF-Pk.

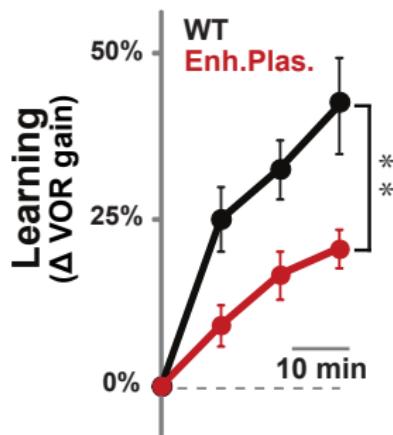
[Marr (1969), Albus (1971), Ito (1972)]

# Enhanced plasticity impairs learning

Knockout of MHC-I K<sup>b</sup>D<sup>b</sup> molecules in PF-Pk synapses

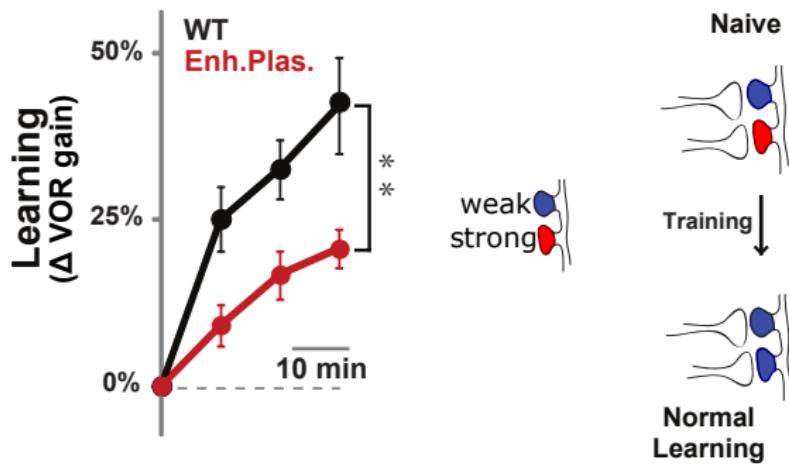
→ lower threshold for LTD → enhanced learning of Rotarod task.

[McConnell et al. (2009)]



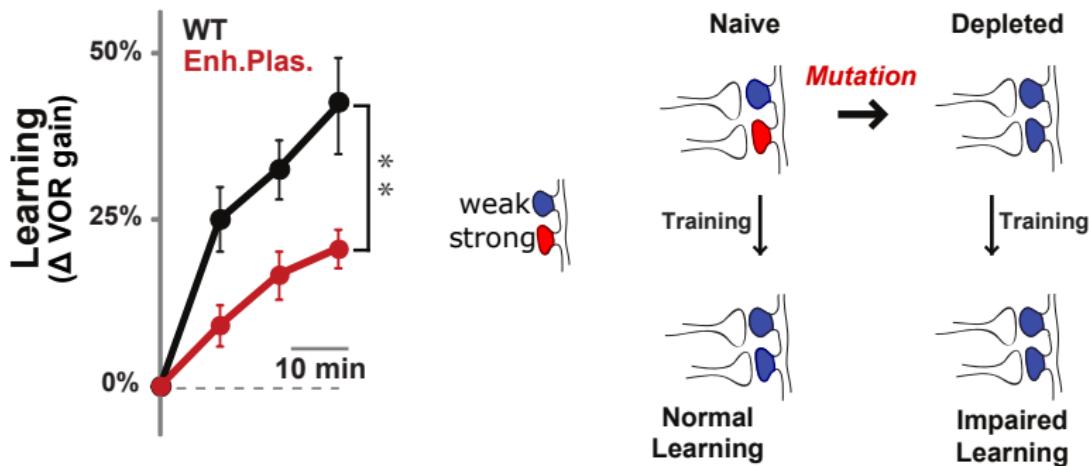
# Enhanced plasticity impairs learning

Learning rate  $\sim$  intrinsic plasticity rate  $\times$  # synapses available for LTD.



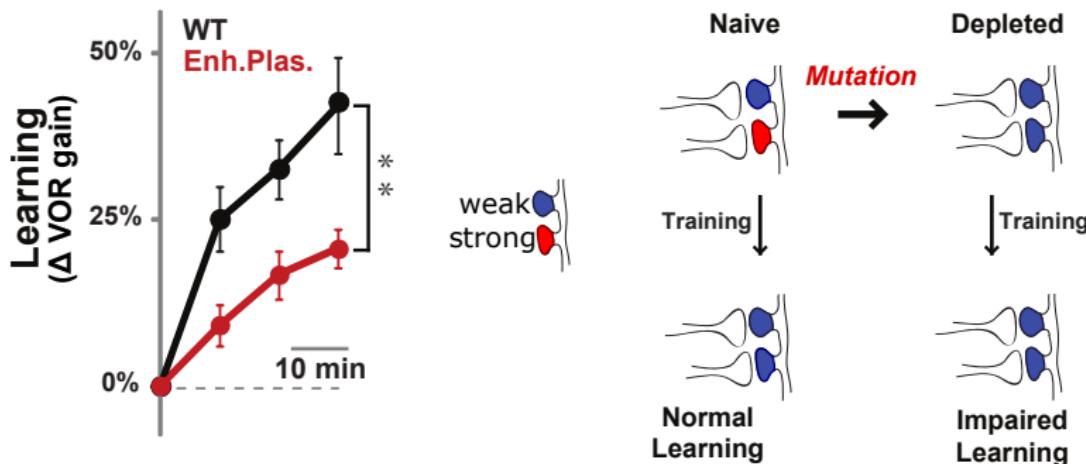
# Enhanced plasticity impairs learning

Learning rate  $\sim$  intrinsic plasticity rate  $\times$  # synapses available for LTD.



# Enhanced plasticity impairs learning

Learning rate  $\sim$  intrinsic plasticity rate  $\times$  # synapses available for LTD.

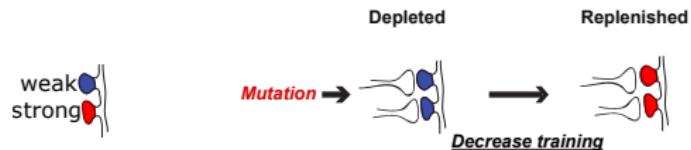


Question 1: depletion effect competes with enhanced intrinsic plasticity.  
When is depletion effect stronger?

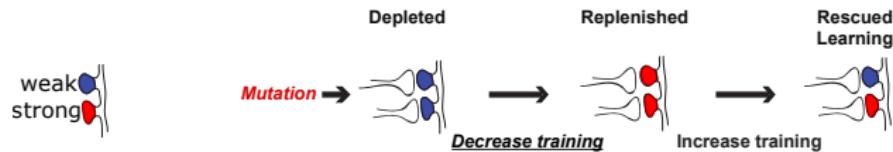
# Replenishment by reverse-training



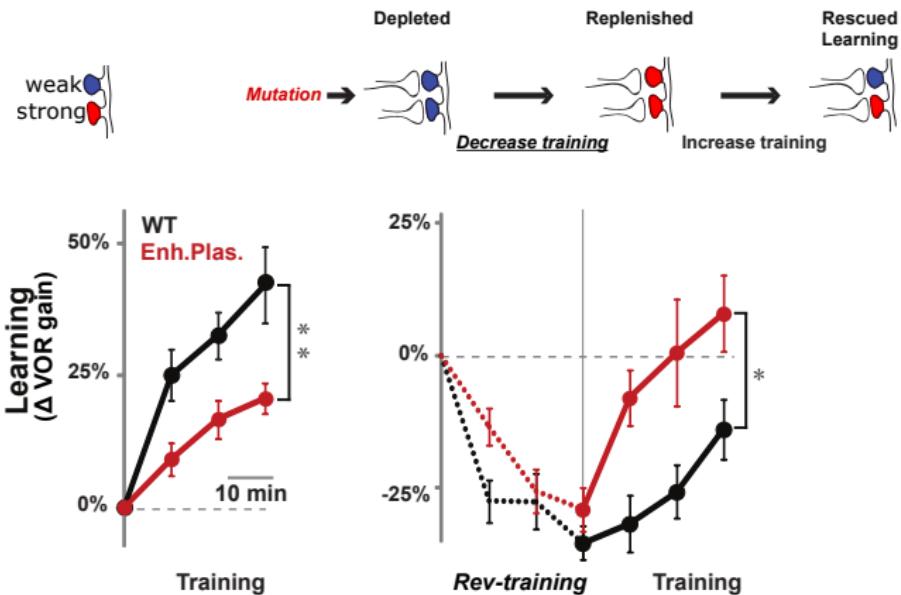
# Replenishment by reverse-training



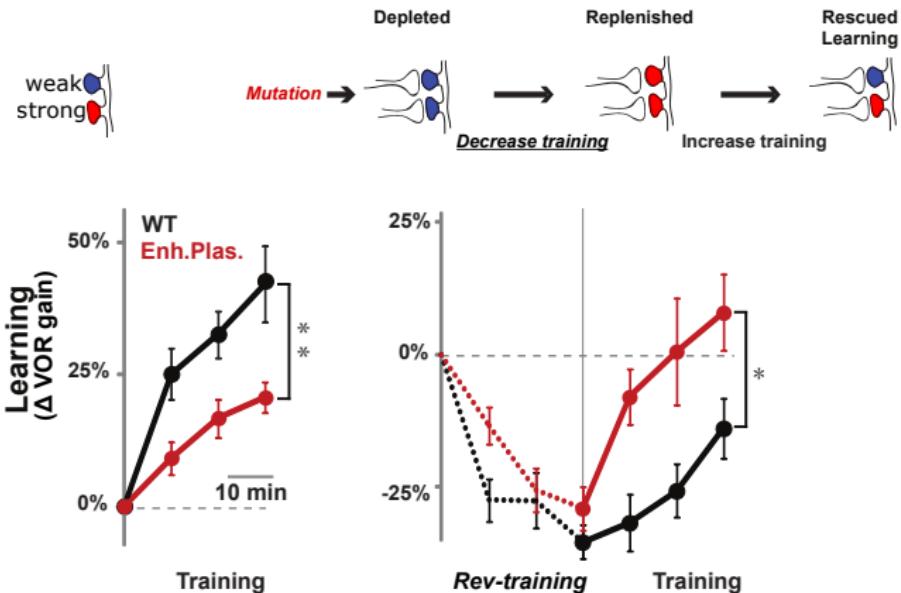
# Replenishment by reverse-training



# Replenishment by reverse-training



# Replenishment by reverse-training

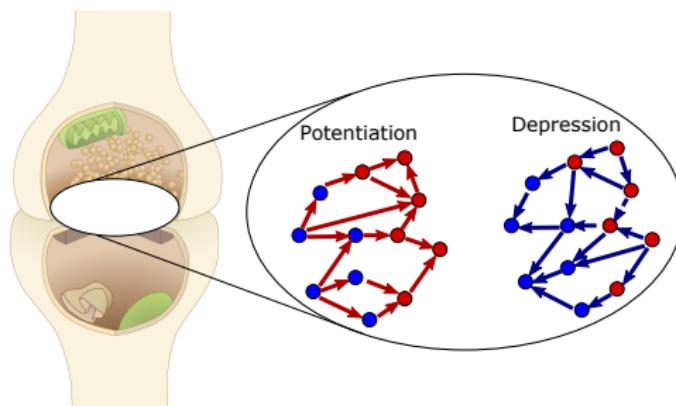


Question 2: When does enhanced plasticity impair learning?

# Models of complex synaptic dynamics

There are  $N$  identical synapses with  $M$  internal functional states.

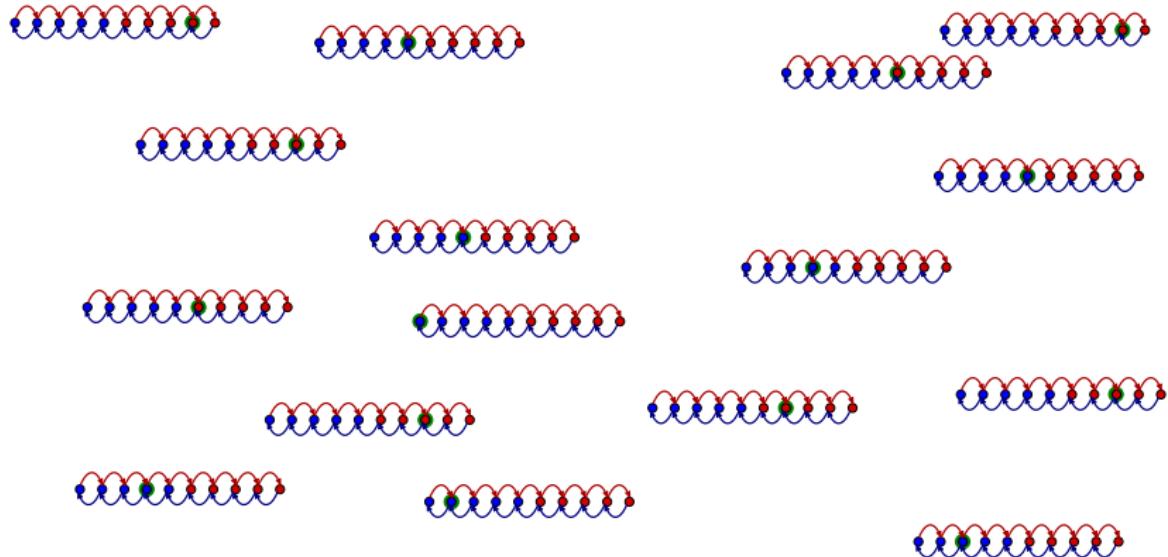
- Internal functional state of synapse → synaptic weight.
  - Candidate plasticity events → transitions between states
- weak  
● strong



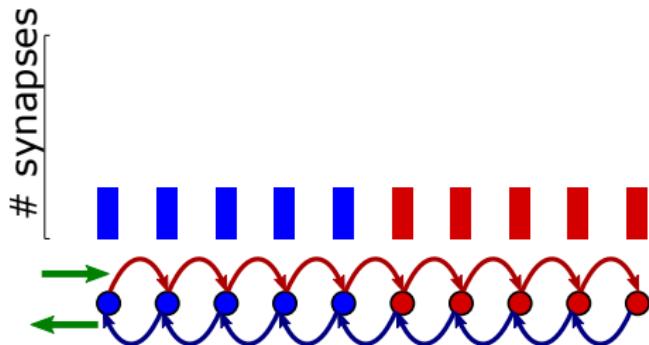
States: #AMPAR, #NMDAR, NMDAR subunit composition,  
CaMK II autophosphorylation, activating PKC, p38 MAPK,...

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

# Modelling VOR experiments

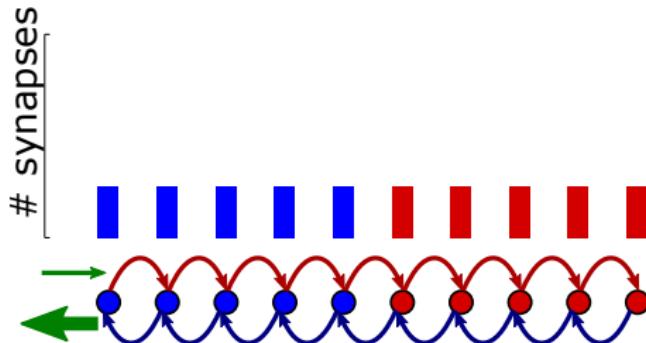


# Modelling VOR experiments



# Modelling VOR experiments

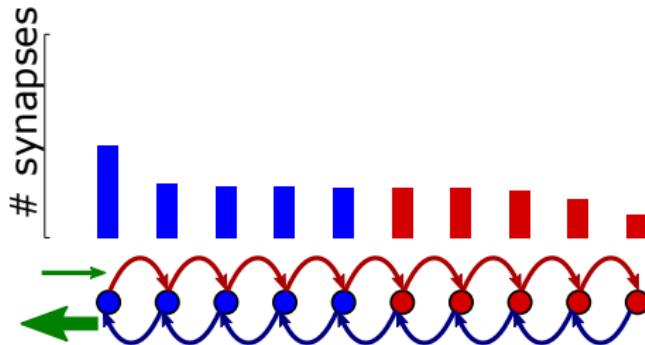
PF-Pk LTD → VOR increase



Training: different CF activity  $\implies$   
change frequency of pot/dep events.

# Modelling VOR experiments

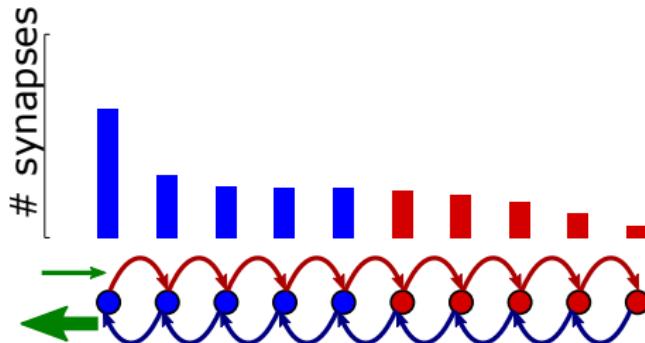
PF-Pk LTD → VOR increase



Training: different CF activity  $\Rightarrow$   
change frequency of pot/dep events.

# Modelling VOR experiments

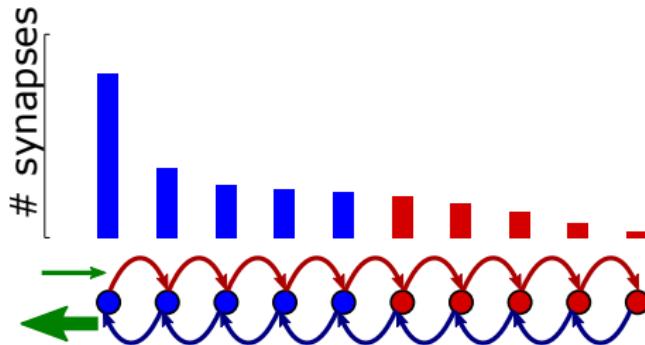
PF-Pk LTD → VOR increase



Training: different CF activity  $\Rightarrow$   
change frequency of pot/dep events.

# Modelling VOR experiments

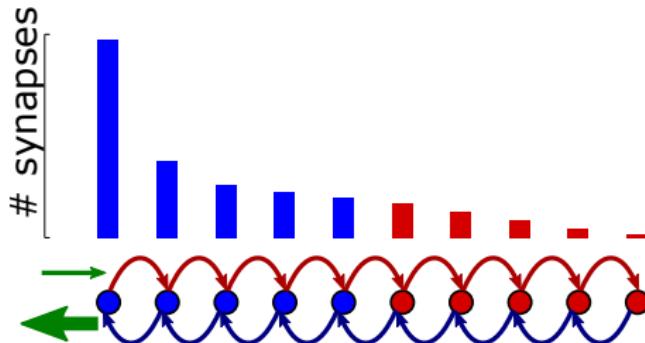
PF-Pk LTD → VOR increase



Training: different CF activity  $\implies$   
change frequency of pot/dep events.

# Modelling VOR experiments

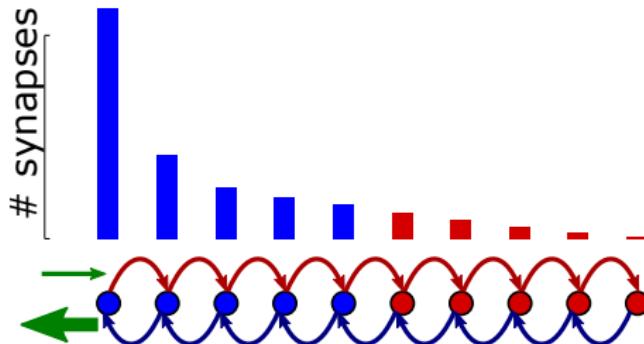
PF-Pk LTD → VOR increase



Training: different CF activity  $\Rightarrow$   
change frequency of pot/dep events.

# Modelling VOR experiments

PF-Pk LTD → VOR increase

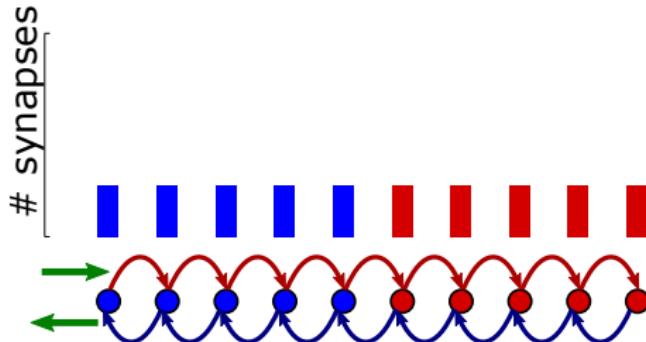


Training: different CF activity  $\implies$   
change frequency of pot/dep events.

Learning: decrease in average synaptic weight.

# Modelling VOR experiments

PF-Pk LTD → VOR increase



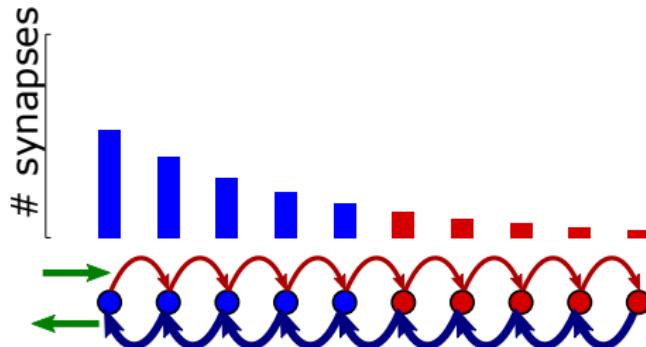
Training: different CF activity  $\Rightarrow$   
change frequency of pot/dep events.

Learning: decrease in average synaptic weight.

Mutation: lower threshold for LTD  $\Rightarrow$   
increase transition probability for depression events.

# Modelling VOR experiments

PF-Pk LTD → VOR increase



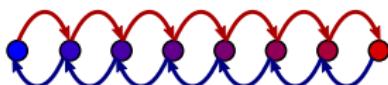
Training: different CF activity  $\Rightarrow$   
change frequency of pot/dep events.

Learning: decrease in average synaptic weight.

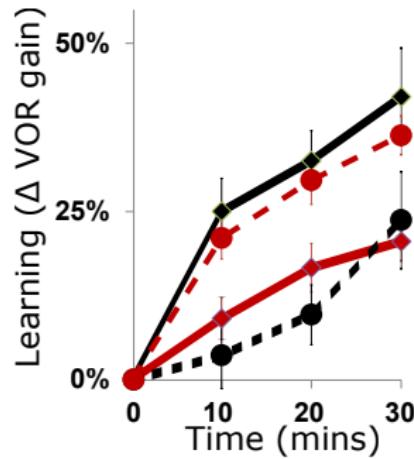
Mutation: lower threshold for LTD  $\Rightarrow$   
increase transition probability for depression events.

# Simple synapses cannot explain the behaviour

Multistate synapse

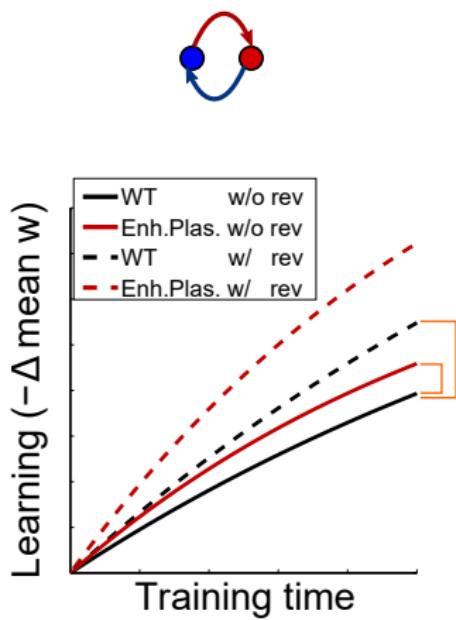


VOR Increase  
Training

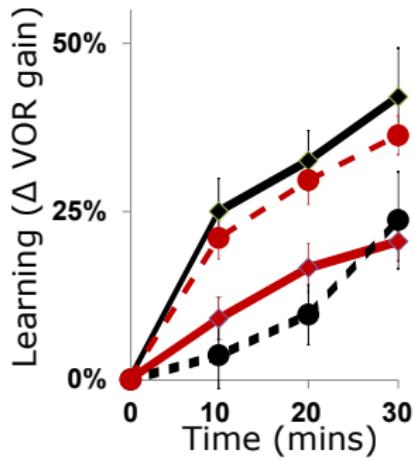


# Simple synapses cannot explain the behaviour

Two-state model

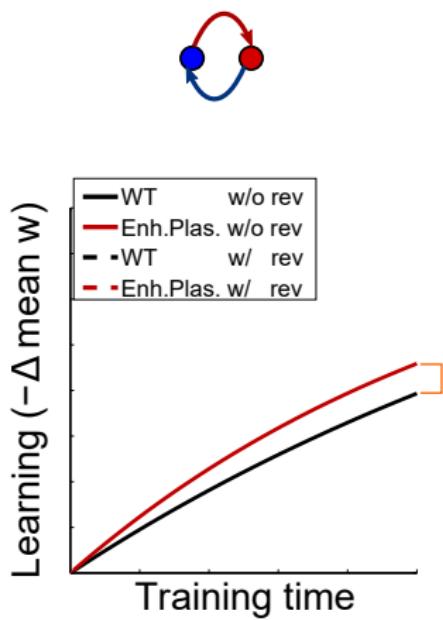


VOR Increase  
Training

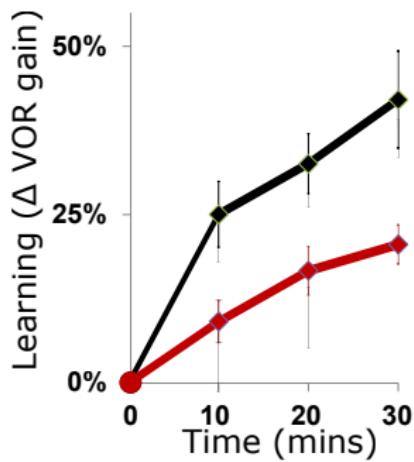


# Simple synapses cannot explain the behaviour

Two-state model



VOR Increase Training

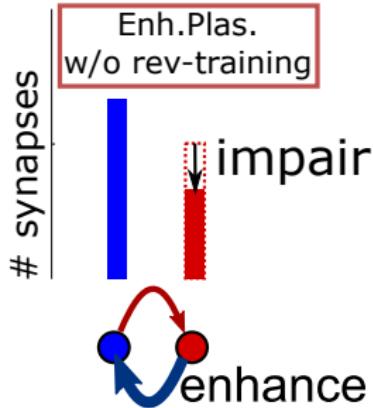


# Simple synapses cannot explain the behaviour

Two-state model



Initial distribution



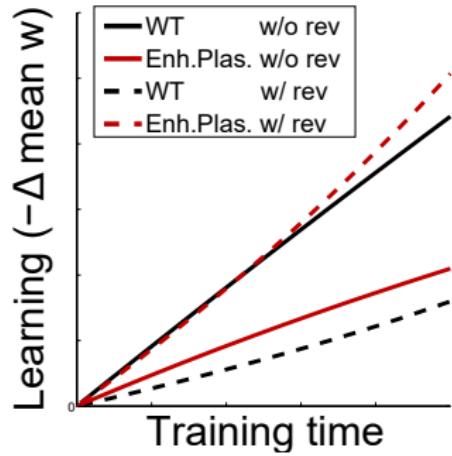
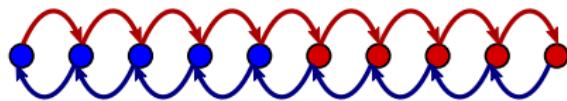
depletion effect

<  
enhanced plasticity

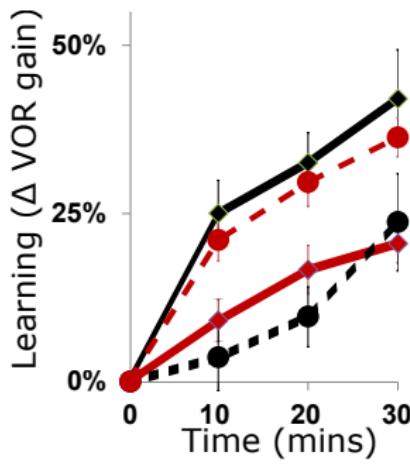
⇒ enhanced learning

# Complex metaplastic synapses can explain the behaviour

Serial model

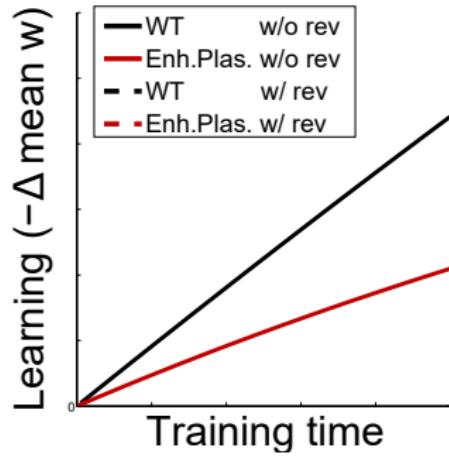
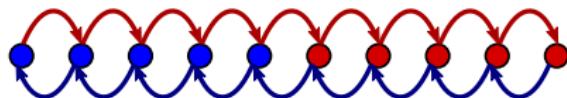


VOR Increase  
Training

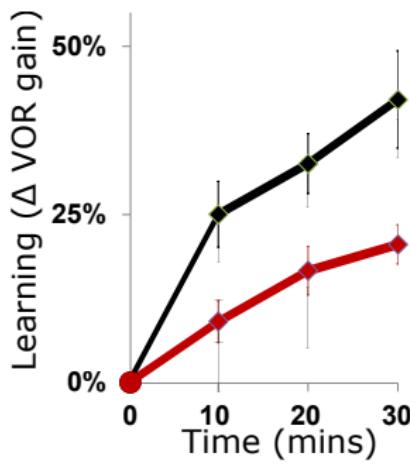


# Complex metaplastic synapses can explain the behaviour

Serial model

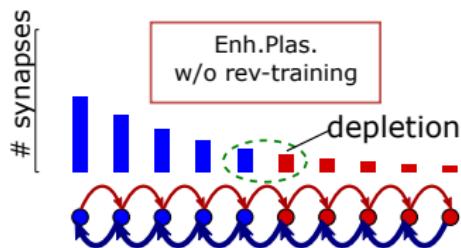
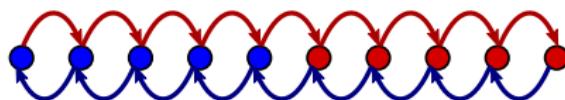


VOR Increase  
Training



# Complex metaplastic synapses can explain the behaviour

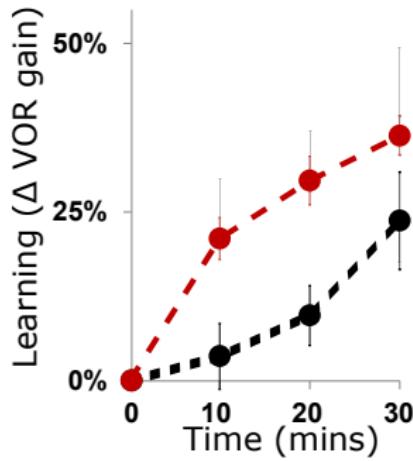
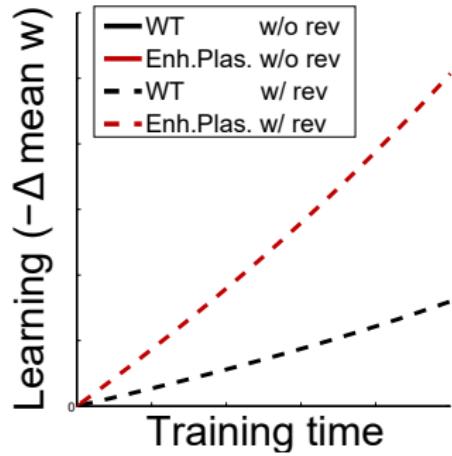
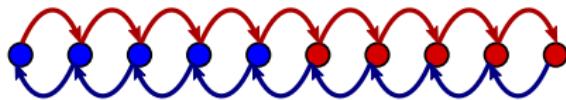
Serial model



amplified depletion  
>  
enhanced plasticity  
⇒ impaired learning

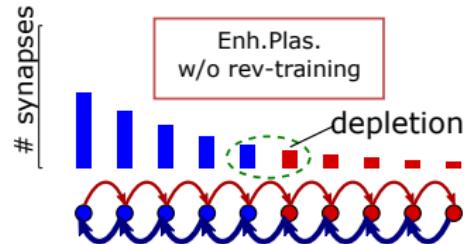
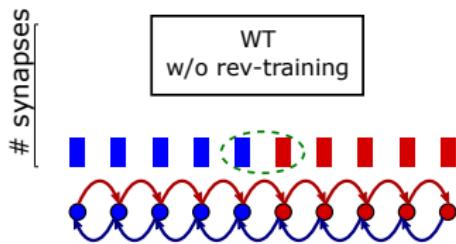
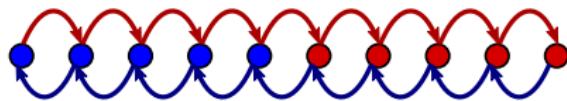
# Complex metaplastic synapses can explain the behaviour

Serial model



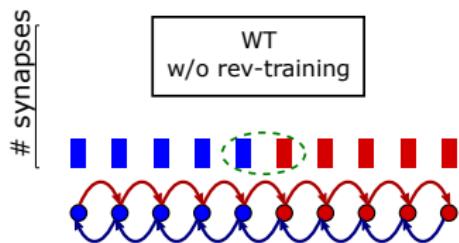
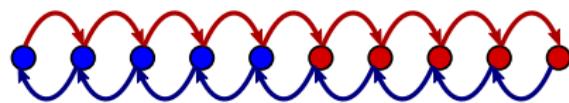
# Complex metaplastic synapses can explain the behaviour

Serial model



Complex metaplastic synapses can explain the behaviour

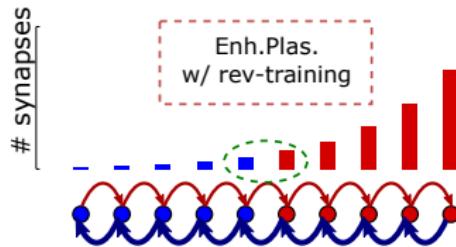
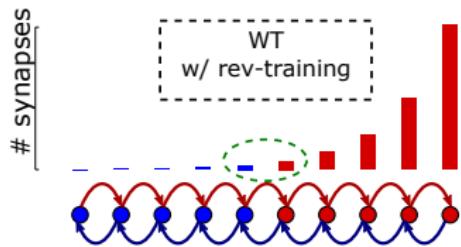
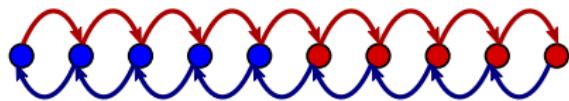
## Serial model



start in: labile states  
↓  
enhanced plasticity  
 $\Rightarrow$  impaired learning

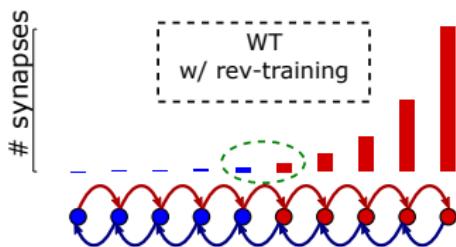
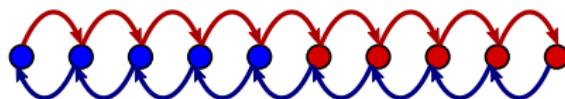
# Complex metaplastic synapses can explain the behaviour

Serial model



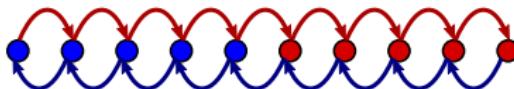
# Complex metaplastic synapses can explain the behaviour

Serial model



start in: stubborn states  
↓  
enhanced plasticity  
⇒ enhanced learning

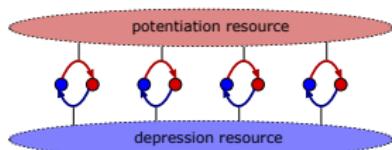
# Essential features of successful models



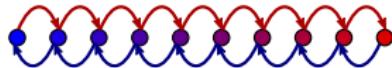
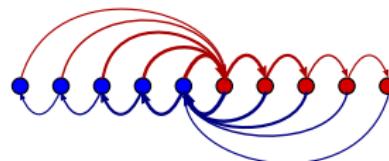
The serial model is successful in reproducing the data due to two features:

- Complexity - needed for depletion to dominate enhanced plasticity,
- Stubbornness - repeated potentiation impairs subsequent depression.

Fail:



Succeed:



# Conclusions

- In VOR learning, depending on prior experience:  
**Enhanced plasticity → enhance/impair learning.**
- **enhanced LTD vs. depletion → learning outcome.**
- Predictions for synaptic physiology:  
**Complexity:** necessary to amplify depletion.  
**Stubbornness:** repeated potentiation impairs subsequent depression.
- We used behaviour to constrain the dynamics of synaptic plasticity.

# Experimental tests?

Traditional experiments:



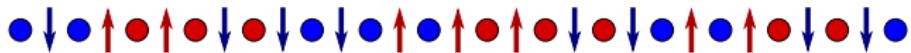
# Experimental tests?

Traditional experiments:

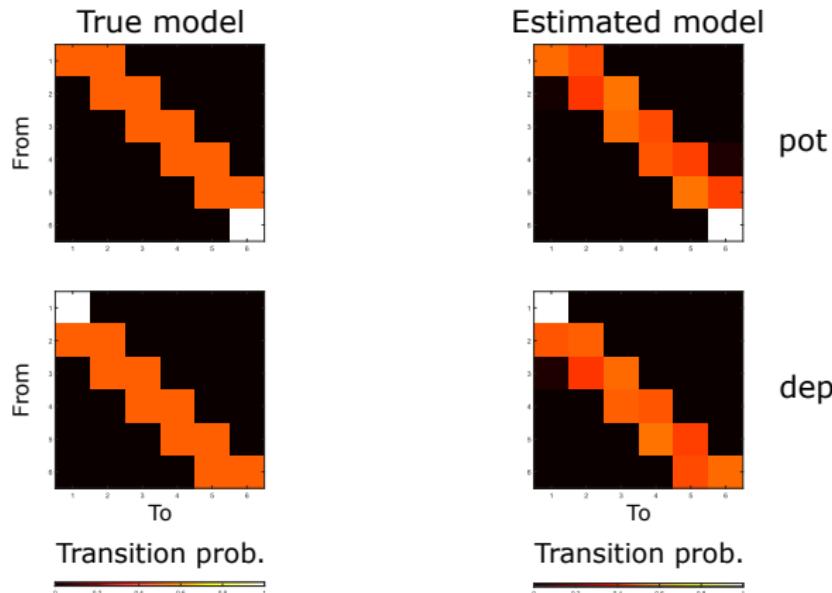


To fit a model: long sequence of small plasticity events.

Observe the changes in synaptic efficacy.



# Simulated experiment



Problem: need *long* sequences.

Whole cell patch of postsynaptic neuron → Ca washout.

# Future directions

## Learning and memory

- Multiple presentations.
- Correlations.
- More realistic tasks.
- Relation to molecular structure?

## Energy use in living systems

- Include space as well as time.
- Coarse graining: molecules → cells → systems.

## High dimensional statistics

- Theory of noisy random projections.

# 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

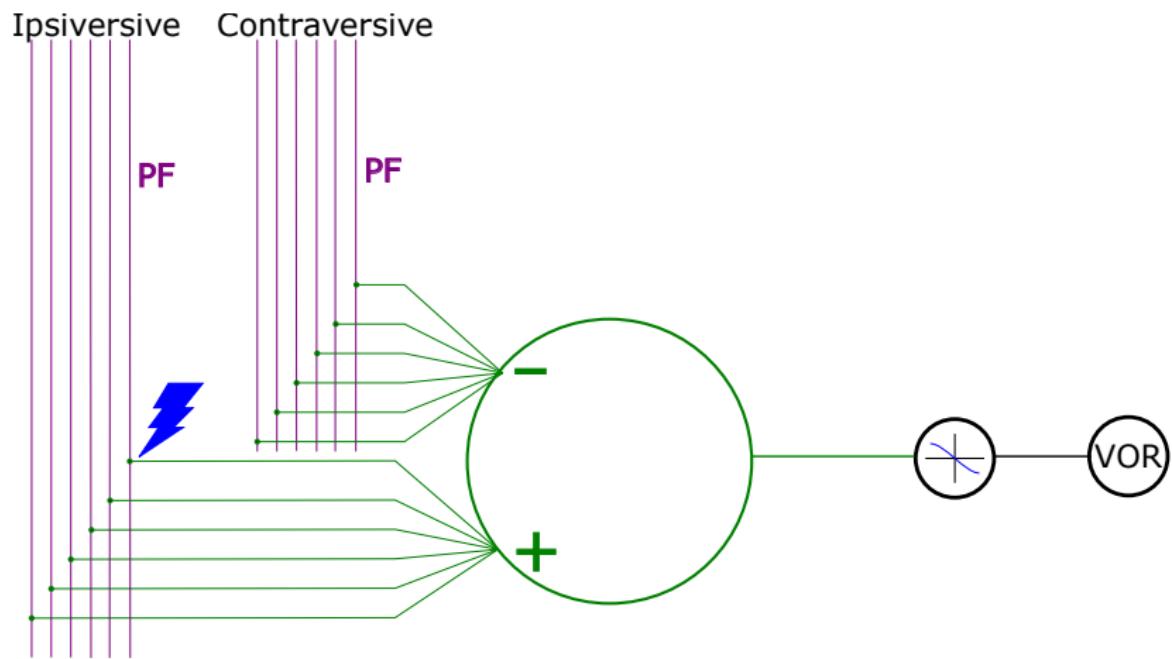
Marcus Benna

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

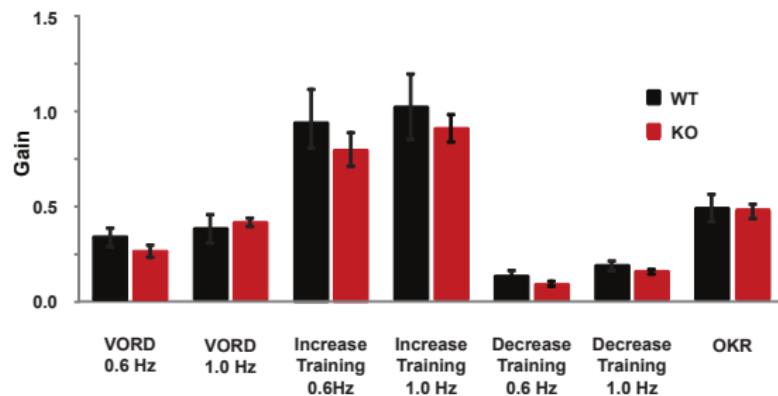
# Summary

- Internal dynamics of synaptic plasticity  
→ understand learning and memory.
- Behaviour → subcellular dynamics of synapses.
- Why & when enhanced plasticity → enhanced/impaired learning.
- Memory envelope: cannot be exceeded by *any* model's memory curve.
- Which synaptic structures are useful for different memory timescales.
- Not just a single model, *all possible models*  
→ general principles relating synaptic structure to function.

# Model of circuit

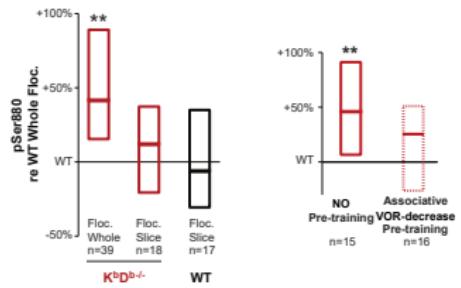


# Baseline



# Evidence: level of depression

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

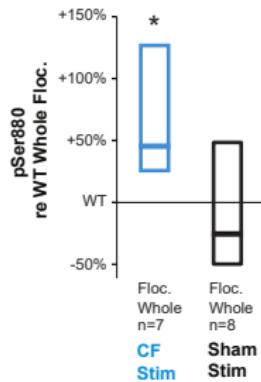
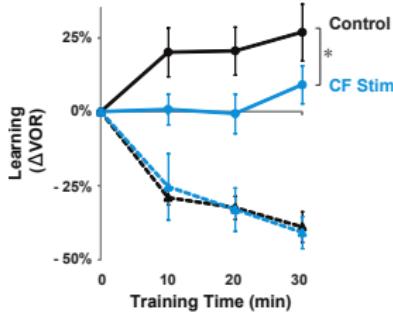
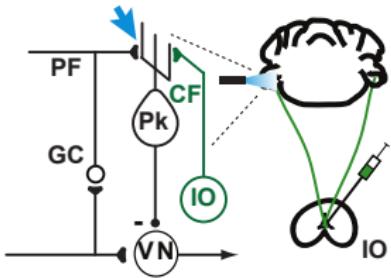


Biochemical signature of PF-Pk LTD.

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

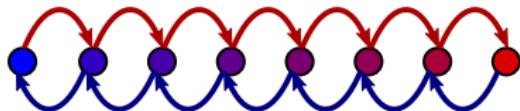
# Evidence: saturation by CF stimulation

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

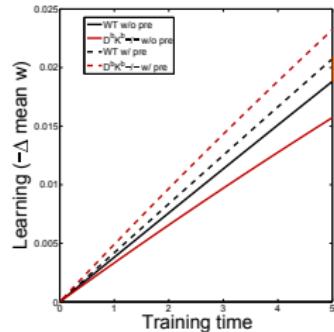
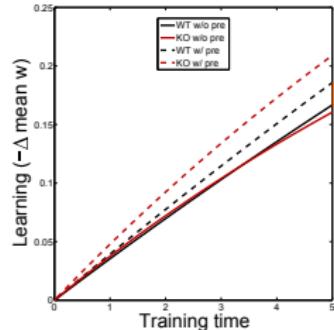
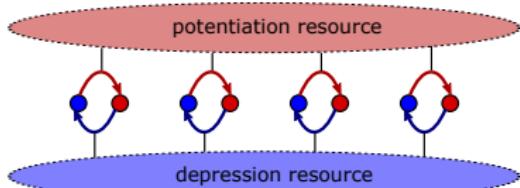


# Other models that fail

## Multistate synapse



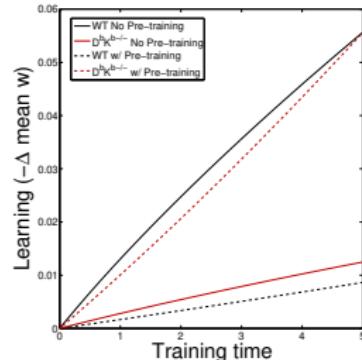
Pooled resource model



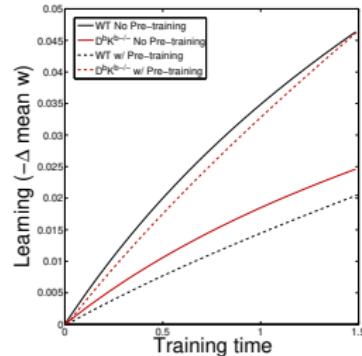
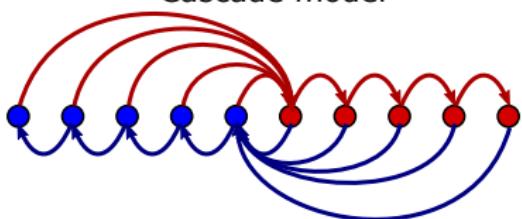
[Amit and Fusi (1994)]

# Other models that work

Non-uniform multistate model

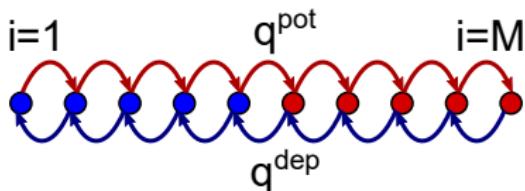


Cascade model



[Fusi et al. (2005)]

## Mathematical explanation



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

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

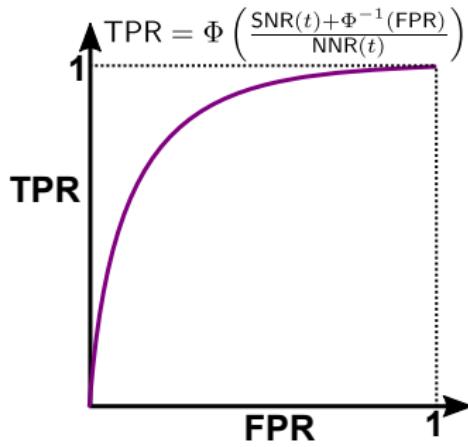
For  $M > 2$ : larger  $q^{\text{dep}}$   $\implies$  slower learning.

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

# Quantifying memory quality

Test if  $\vec{w}_{\text{ideal}} \cdot \vec{w}(t) \geq \theta?$

[Sommer and Dayan (1998)]



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

$$\overline{\text{SNR}}(\tau) = \int d\tau \frac{e^{-t/\tau}}{\tau} \text{SNR}(t).$$

$$\text{NNR}(t) = \sqrt{\frac{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(t))}{\text{Var}(\vec{w}_{\text{ideal}} \cdot \vec{w}(\infty))}}.$$

Also: KL divergence, Chernoff distance, . . .

# Parameters for synaptic dynamics

$f^{\text{pot/dep}}$  = fraction of events that are pot/dep,

pot. event:  $M_{ij}^{\text{pot}}$  = transition prob.  $i \rightarrow j$ ,

$$\mathbf{W}^{\text{pot}} = f^{\text{pot}}(\mathbf{M}^{\text{pot}} - \mathbf{I}),$$

dep. event:  $M_{ij}^{\text{dep}}$  = transition prob.  $i \rightarrow j$ ,

$$\mathbf{W}^{\text{dep}} = f^{\text{dep}}(\mathbf{M}^{\text{dep}} - \mathbf{I}).$$

Constraints:

$$f^{\text{pot/dep}}, M_{ij}^{\text{pot/dep}} \in [0, 1], \quad f^{\text{pot}} + f^{\text{dep}} = \sum_j M_{ij}^{\text{pot/dep}} = 1.$$

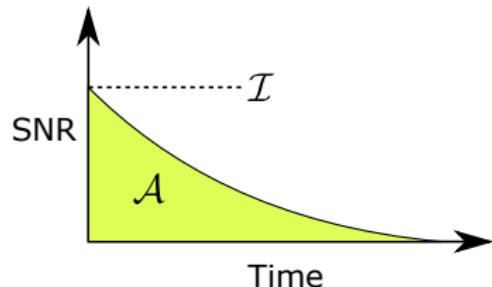
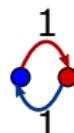
Memory curve given by

$$\begin{aligned} \overline{\text{SNR}}(\tau) &= \sqrt{N} \pi (\mathbf{W}^{\text{pot}} - \mathbf{W}^{\text{dep}}) \left[ \mathbf{I} - r\tau (\mathbf{W}^{\text{pot}} + \mathbf{W}^{\text{dep}}) \right]^{-1} \mathbf{w}. \\ &= \sqrt{N} \sum_a \frac{\mathcal{I}_a}{1 + r\tau/\tau_a}. \end{aligned}$$

# Upper bounds on measures of memory

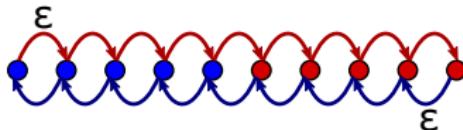
Initial SNR:

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



Area under curve:

$$\mathcal{A} = \int_0^\infty \text{SNR}(t) dt = \sum_a \mathcal{I}_a \tau_a \leq \sqrt{N}(M-1)/r.$$



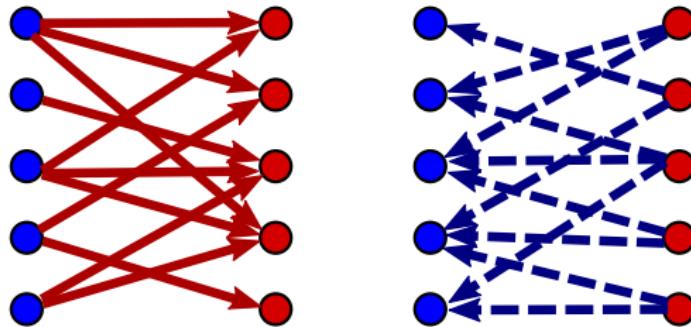
[Lahiri and Ganguli (2013)]

## Initial SNR as flux

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

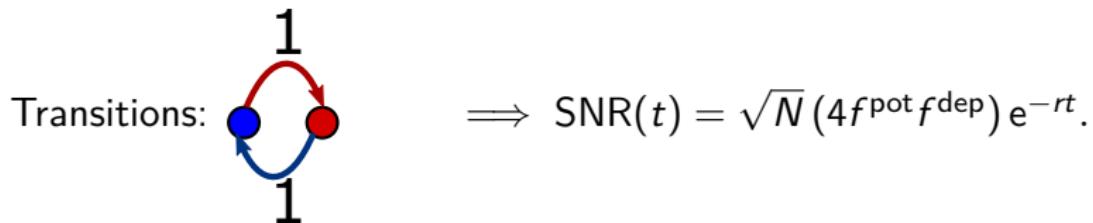
$$\text{SNR}(0) \leq \frac{4\sqrt{N}}{r} \Phi_{-+}.$$

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



## Two-state model

Two-state model equivalent to previous slide:



Maximal initial SNR:

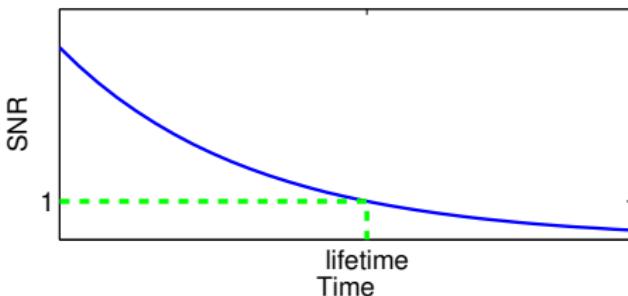
$$\text{SNR}(0) \leq \sqrt{N}.$$

## Area under memory curve

$$\mathcal{A} = \int_0^\infty dt \text{ SNR}(t), \quad \overline{\text{SNR}}(\tau) \rightarrow \frac{\mathcal{A}}{\tau} \quad \text{as} \quad \tau \rightarrow \infty.$$

Area bounds memory lifetime:

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



This area has an upper bound:

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

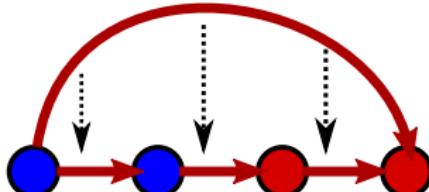
Saturated by a model with linear chain topology.

## 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

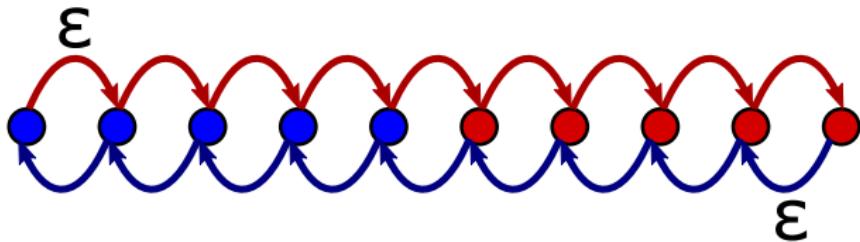
$$A = \frac{2\sqrt{N}}{r} \sum_k \pi_k |k - \langle k \rangle|.$$

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.

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

## Technical detail: ordering states

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

$$\eta = \sum_j \mathbf{T}_{ij} \pi_j,$$

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

[Kemeny and Snell (1960)]

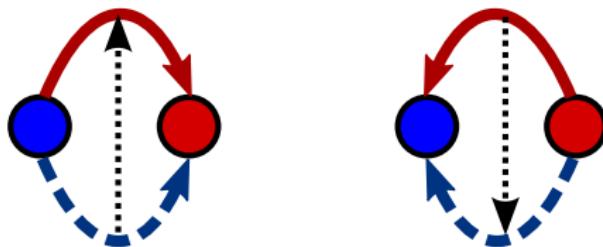
We define:

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

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

## Technical detail: upper/lower triangular

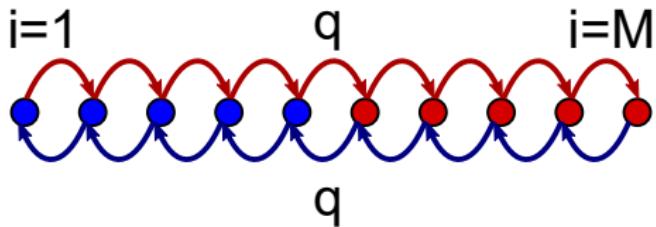
With states in order:



Endpoint: potentiation goes right, depression goes left.

back

## Intuition for using topology



$$\begin{array}{ll} \mathcal{I} \propto q, & \max_a \tau_a \propto \frac{1}{q}, \\ \mathcal{I} \propto \frac{1}{M}, & \max_a \tau_a \propto M^2, \end{array} \implies \begin{array}{ll} \text{Stochasticity: } \mathcal{I} \propto \frac{1}{\tau_{\max}}, & \\ \text{Topology: } \mathcal{I} \propto \frac{1}{\sqrt{\tau_{\max}}}. & \end{array}$$

# 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 4 5



D. J. Amit and S. Fusi.

“Constraints on learning in dynamic synapses”.

*Network: Comp. Neural*, 3(4):443–464, (1992) .

6



D. J. Amit and S. Fusi.

“Learning in neural networks with material synapses”.

*Neural Comput.*, 6(5):957–982, (1994) .

6 101

## 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) .

6



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) .

6

## 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) .

7 8 9



Johanna M. Montgomery and Daniel V. Madison.

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

*Neuron*, 33(5):765 – 777, (2002) .

7 8 9

# References IV



S. Fusi, P. J. Drew, and L. F. Abbott.

“Cascade models of synaptically stored memories”.

*Neuron*, 45(4):599–611, (February, 2005) .

8 9 14 37 66 102



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].

8 9



S. Fusi and L. F. Abbott.

“Limits on the memory storage capacity of bounded synapses”.

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

14 66

## References V

 A. B. Barrett and M. C. van Rossum.  
“Optimal learning rules for discrete synapses”.

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

14 66 110

 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) .

14

# 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>.

14

106



Christian Leibold and Richard Kempler.

“Sparseness Constrains the Prolongation of Memory Lifetime via Synaptic Metaplasticity”.

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

37

# References VII

 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) .

37

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

“Evolution of Synapse Complexity and Diversity”.  
*Annual Review of Neuroscience*, 35(1):111–131, (2012) .

44

45

 Larry R Squire and Pablo Alvarez.

“Retrograde amnesia and memory consolidation: a neurobiological perspective”.

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

44

45

## References VIII



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.

44

45



Jan Born and Ines Wilhelm.

"System consolidation of memory during sleep.".

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

44

45



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) .

44

45

# References IX



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) .

44

45



Y. P. Tang, E. Shimizu, G. R. Dube, C. Rampon, G. A. Kerchner, M. Zhuo, G. Liu, and J. Z. Tsien.

“Genetic enhancement of learning and memory in mice”.

*Nature*, 401(6748):63–69, (Sep, 1999) .

52

## References X



Gaël Malleret, Ursula Haditsch, David Genoux, Matthew W. Jones, Tim V.P. Bliss, Amanda M. Vanhoose, Carl Weitlauf, Eric R. Kandel, Danny G. Winder, and Isabelle M. Mansuy.

"Inducible and Reversible Enhancement of Learning, Memory, and Long-Term Potentiation by Genetic Inhibition of Calcineurin".

*Cell*, 104(5):675 – 686, (2001) .

52



J. S. Guan, S. J. Haggarty, E. Giacometti, J. H. Dannenberg, N. Joseph, J. Gao, T. J. Nieland, Y. Zhou, X. Wang, R. Mazitschek, J. E. Bradner, R. A. DePinho, R. Jaenisch, and L. H. Tsai.

"HDAC2 negatively regulates memory formation and synaptic plasticity".

*Nature*, 459(7243):55–60, (May, 2009) .

52

## References XI



M. Migaud, P. Charlesworth, M. Dempster, L. C. Webster, A. M. Watabe, M. Makhinson, Y. He, M. F. Ramsay, R. G. Morris, J. H. Morrison, T. J. O'Dell, and S. G. Grant.

"Enhanced long-term potentiation and impaired learning in mice with mutant postsynaptic density-95 protein".

*Nature*, 396(6710):433–439, (December, 1998) .

53



N. Uetani, K. Kato, H. Ogura, K. Mizuno, K. Kawano, K. Mikoshiba, H. Yakura, M. Asano, and Y. Iwakura.

"Impaired learning with enhanced hippocampal long-term potentiation in PTPdelta-deficient mice".

*EMBO J.*, 19(12):2775–2785, (June, 2000) .

53

## References XII



Patrick R Cox, Velia Fowler, Bisong Xu, J.David Sweatt, Richard Paylor, and Huda Y Zoghbi.

“Mice lacking tropomodulin-2 show enhanced long-term potentiation, hyperactivity, and deficits in learning and memory”.

*Molecular and Cellular Neuroscience*, 23(1):1 – 12, (2003) .

53



S.K.E. Koekkoek, K. Yamaguchi, B.A. Milojkovic, B.R. Dortland, T.J.H. Ruigrok, R. Maex, W. De Graaf, A.E. Smit, F. VanderWerf, C.E. Bakker, R. Willemsen, T. Ikeda, S. Kakizawa, K. Onodera, D.L. Nelson, E. Mientjes, M. Joosten, E. De Schutter, B.A. Oostra, M. Ito, and C.I. De Zeeuw.

“Deletion of *FMR1* in Purkinje Cells Enhances Parallel Fiber LTD, Enlarges Spines, and Attenuates Cerebellar Eyelid Conditioning in Fragile X Syndrome”.

*Neuron*, 47(3):339 – 352, (2005) .

53

## References XIII



Mansuo L Hayashi, Se-Young Choi, B.S.Shankaranarayana Rao, Hae-Yoon Jung, Hey-Kyoung Lee, Dawei Zhang, Sumantra Chattarji, Alfredo Kirkwood, and Susumu Tonegawa.

"Altered Cortical Synaptic Morphology and Impaired Memory Consolidation in Forebrain- Specific Dominant-Negative {PAK} Transgenic Mice".

*Neuron*, 42(5):773 – 787, (2004) .

53



Kris Rutten, Dinah L. Misner, Melissa Works, Arjan Blokland, Thomas J. Novak, Luca Santarelli, and Tanya L. Wallace.

"Enhanced long-term potentiation and impaired learning in phosphodiesterase 4D-knockout (PDE4D-/-) mice".

*European Journal of Neuroscience*, 28(3):625–632, (2008) .

53

# References XIV



David Marr.

“A theory of cerebellar cortex”.

*The Journal of Physiology*, 202(2):437–470, (1969) .

56



James S. Albus.

“A theory of cerebellar function”.

*Mathematical Biosciences*, 10(1):25 – 61, (1971) .

56



Masao Ito.

“Neural design of the cerebellar motor control system”.

*Brain research*, 40(1):81–84, (1972) .

56

## References XV

-  Michael J. McConnell, Yanhua H. Huang, Akash Datwani, and Carla J. Shatz.  
“H2-Kb and H2-Db regulate cerebellar long-term depression and limit motor learning”.  
*Proc. Natl. Acad. Sci. U.S.A.*, 106(16):6784–6789, (2009) .  
57
-  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) .  
104
-  J.G. Kemeny and J.L. Snell.  
*Finite markov chains*.  
Springer, 1960.  
112