Microsoft Research

Each year Microsoft Research hosts hundreds of influential speakers from around the world including leading scientists, renowned experts in technology, book authors, and leading academics, and makes videos of these lectures freely available.

NIPS Thanks Its Sponsors

Air Force Office of Scientific Research
amazon.com
Microsoft Research
Google
facebook
SKYTREE
THE MACHINE LEARNING COMPANY
TWO SIGMA
United Technologies Research Center
D E Shaw & Co
xerox
IBM Research
YAHOO! LABS
DRW TRADING GROUP
TOYOTA
millionshort
criteo
PDT PARTNERS
Springer
Machine Learning Journal
Disney Research

Hippocampal memory reactivation in awake and sleep states

Matthew Wilson
Departments of Brain and Cognitive Sciences and Biology
MIT

The Picower Institute
for learning and memory

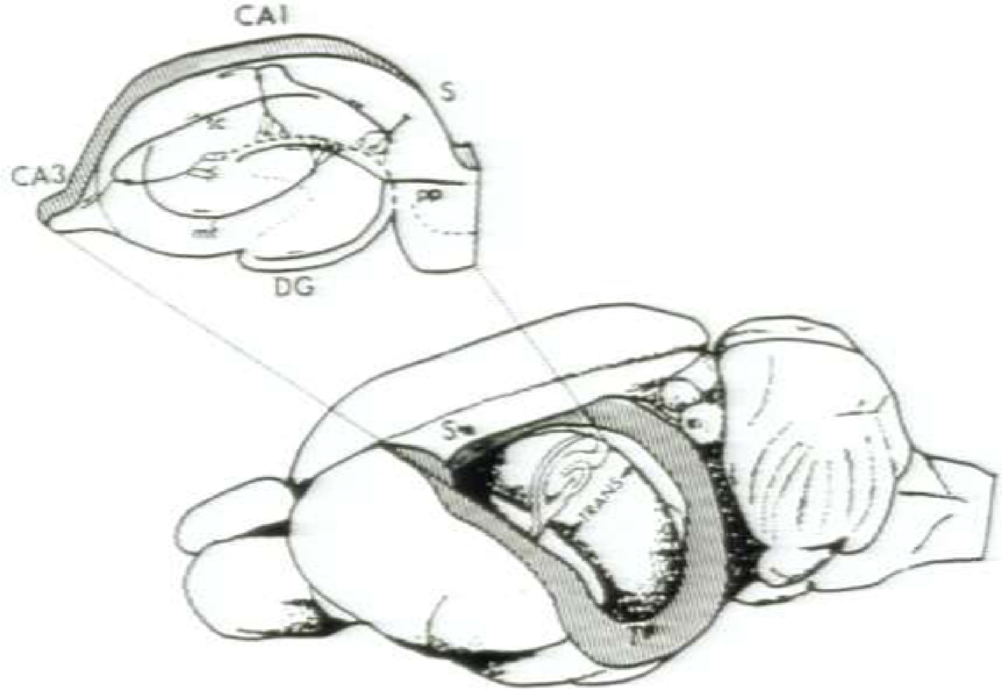


Fig. 2. The position of the hippocampal formation in the rat brain is shown in this drawing of a preparation in which the cortical surface overlying the hippocampus has been removed. The hippocampus is an elongated, C-shaped structure with the long or septotemporal axis running from the septal nuclei rostrally (S) to the temporal cortex (T) ventrocaudally. The short or transverse axis (TRANS) is oriented perpendicular to the septotemporal axis. The major fields of the hippocampal formation (except for the entorhinal cortex) are found in slices taken approximately midway along the septotemporal axis. The slice pictured at top left is a representation of the summary of the major neuronal elements and intrinsic connections of the hippocampal formation as originally illustrated by Andersen *et al.* (see text for details).

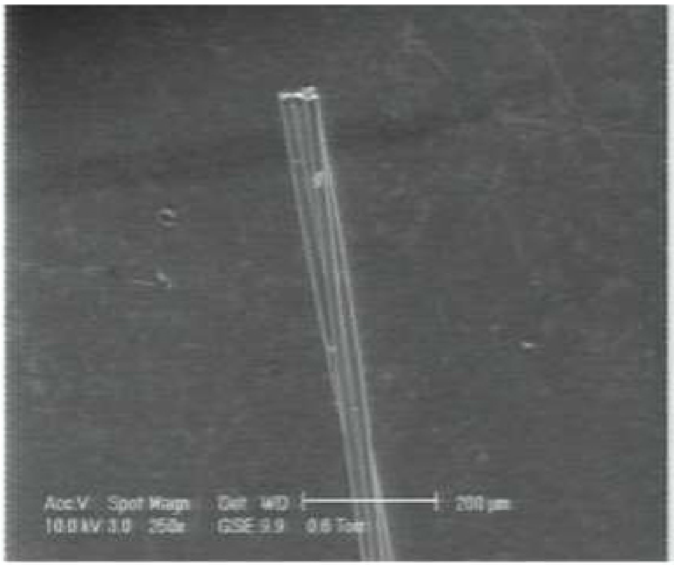
Abbreviations: DG, dentate gyrus; mf, mossy fibers; pp, perforant path; sc, Schaffer collaterals.

From Amaral and Witter, 1989

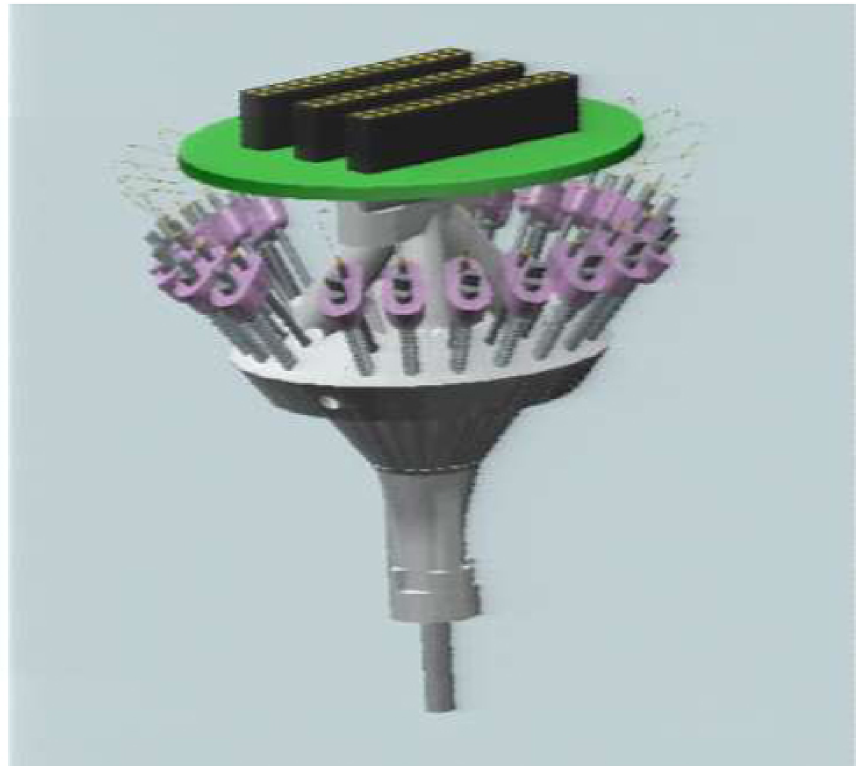
Hippocampus in spatial and episodic memory

- The hippocampus is involved in the formation of episodic memory as well as spatial memory used in navigation.
- Navigation - linkage of spatial locations
- Episodic memory - linkage of events
- Both may depend critically on temporal sequence encoding

Neural recording device

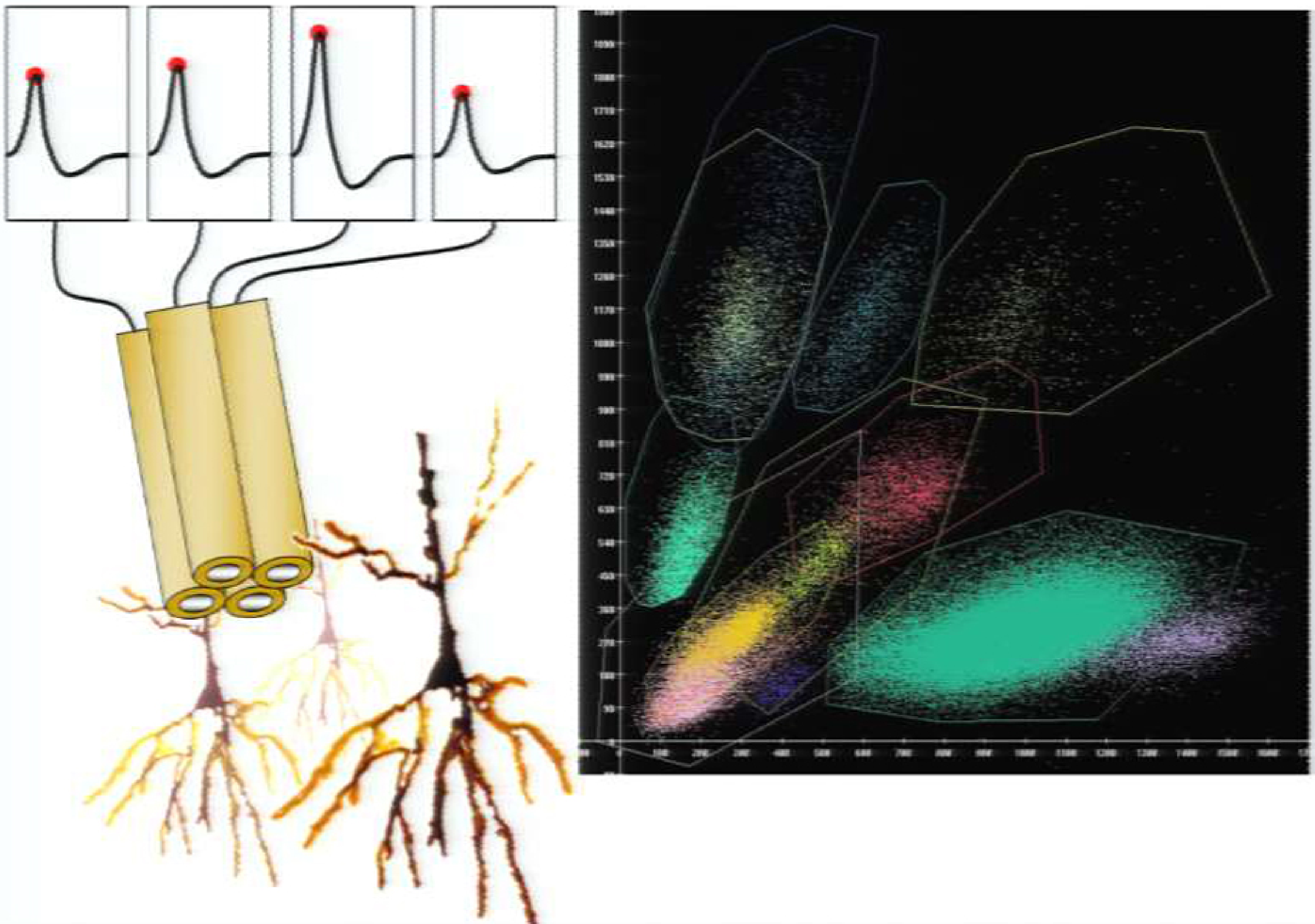


4-channel microwire electrode

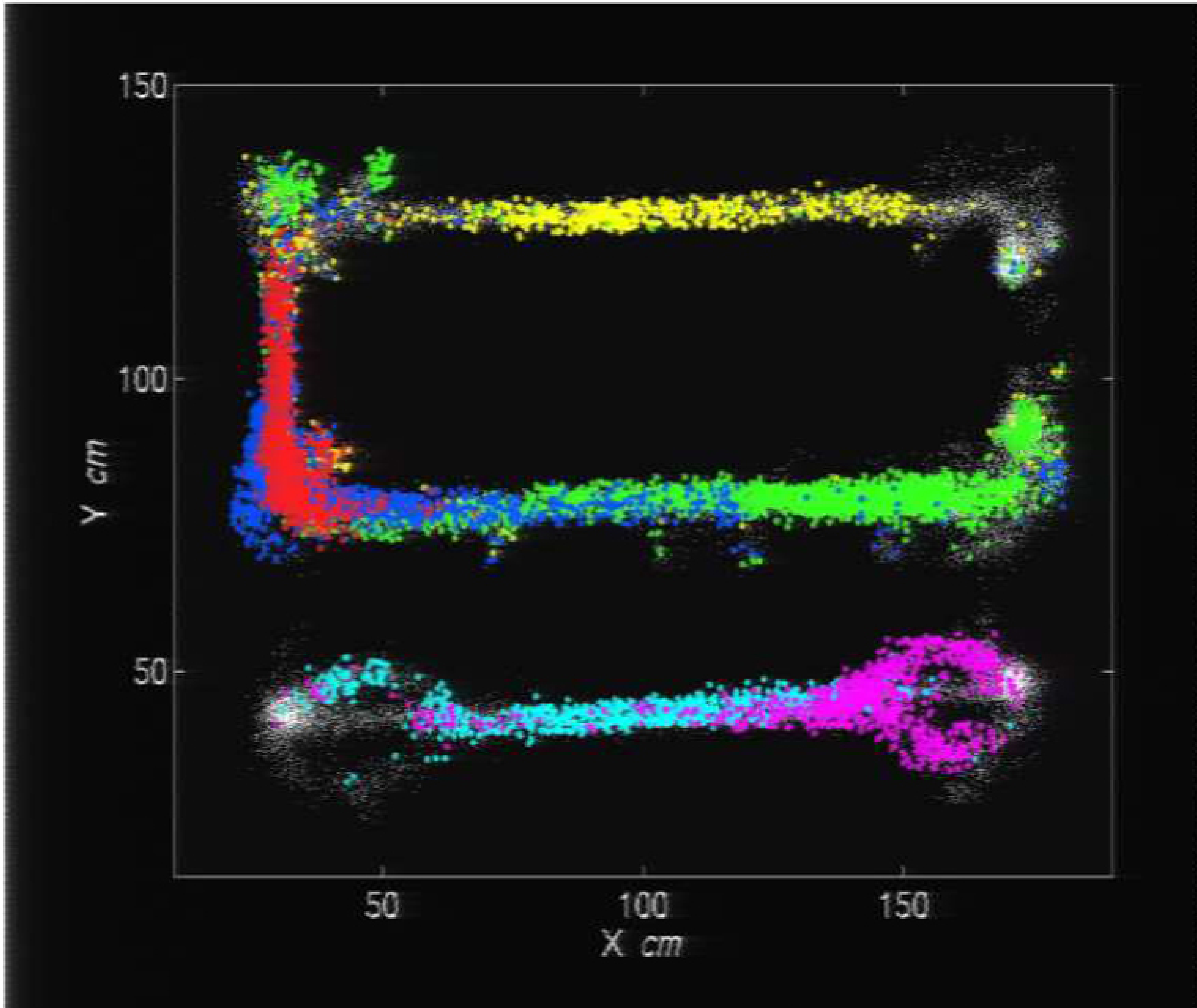


Multiple electrode microdrive array

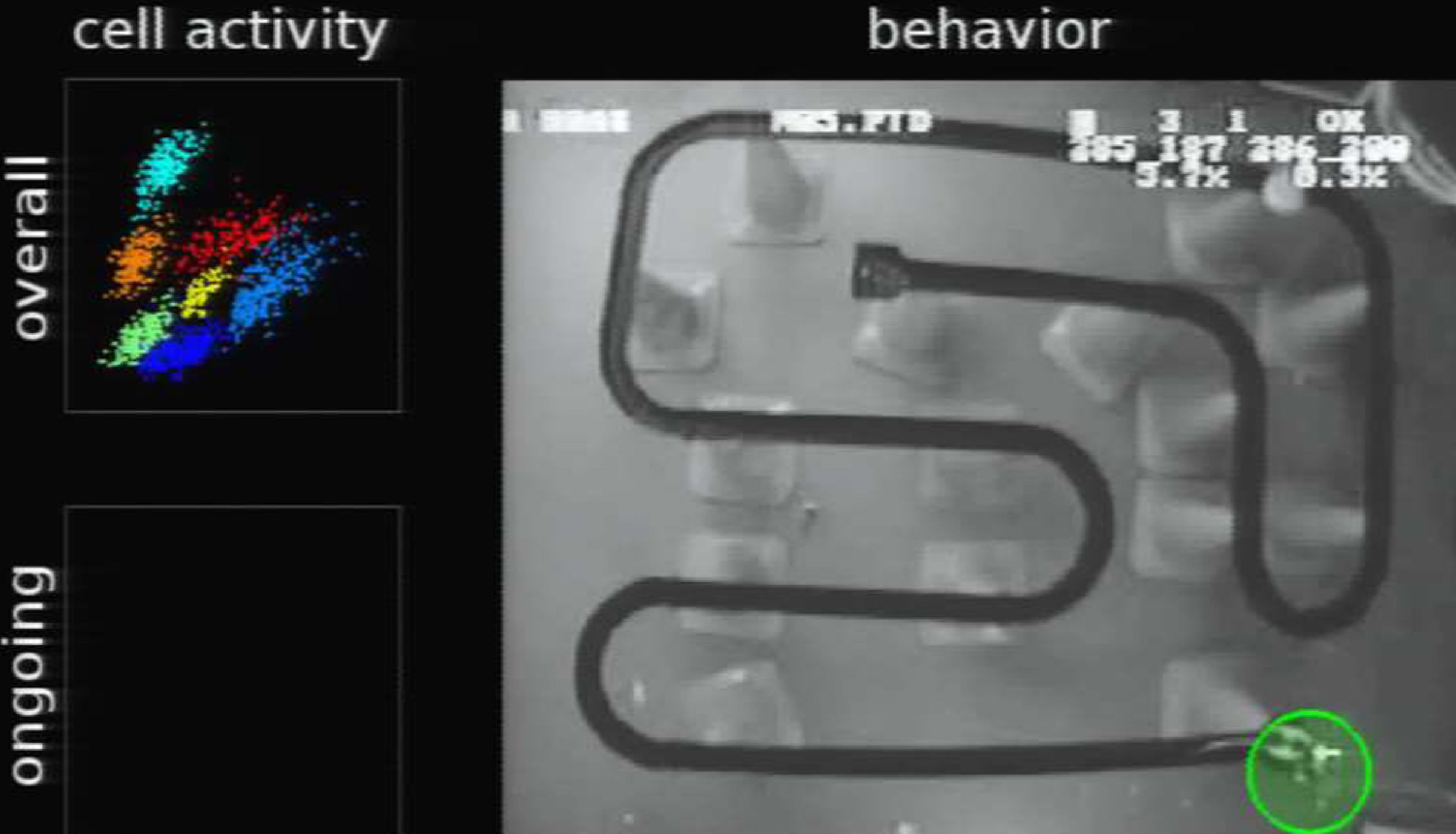
Spike amplitude clustering



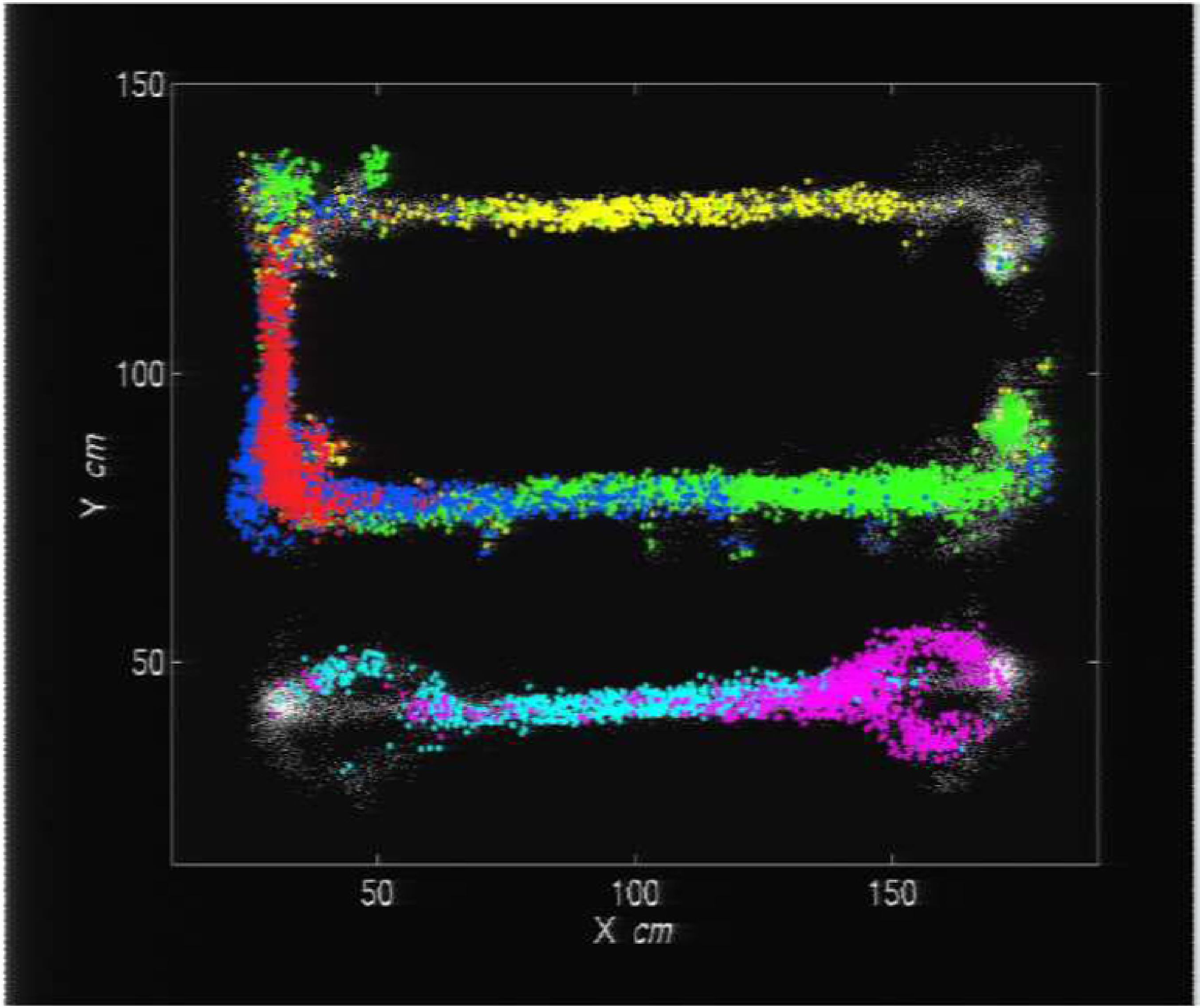
Place Fields on Linear Tracks



Hippocampal Place Cells



Place Fields on Linear Tracks

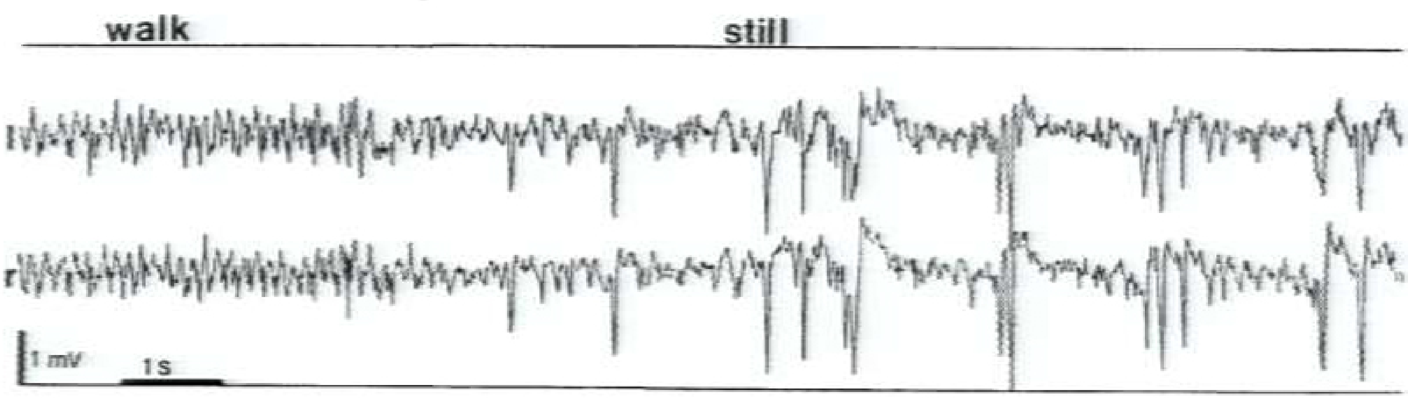


Hippocampal Place Cells

Hippocampus online and offline

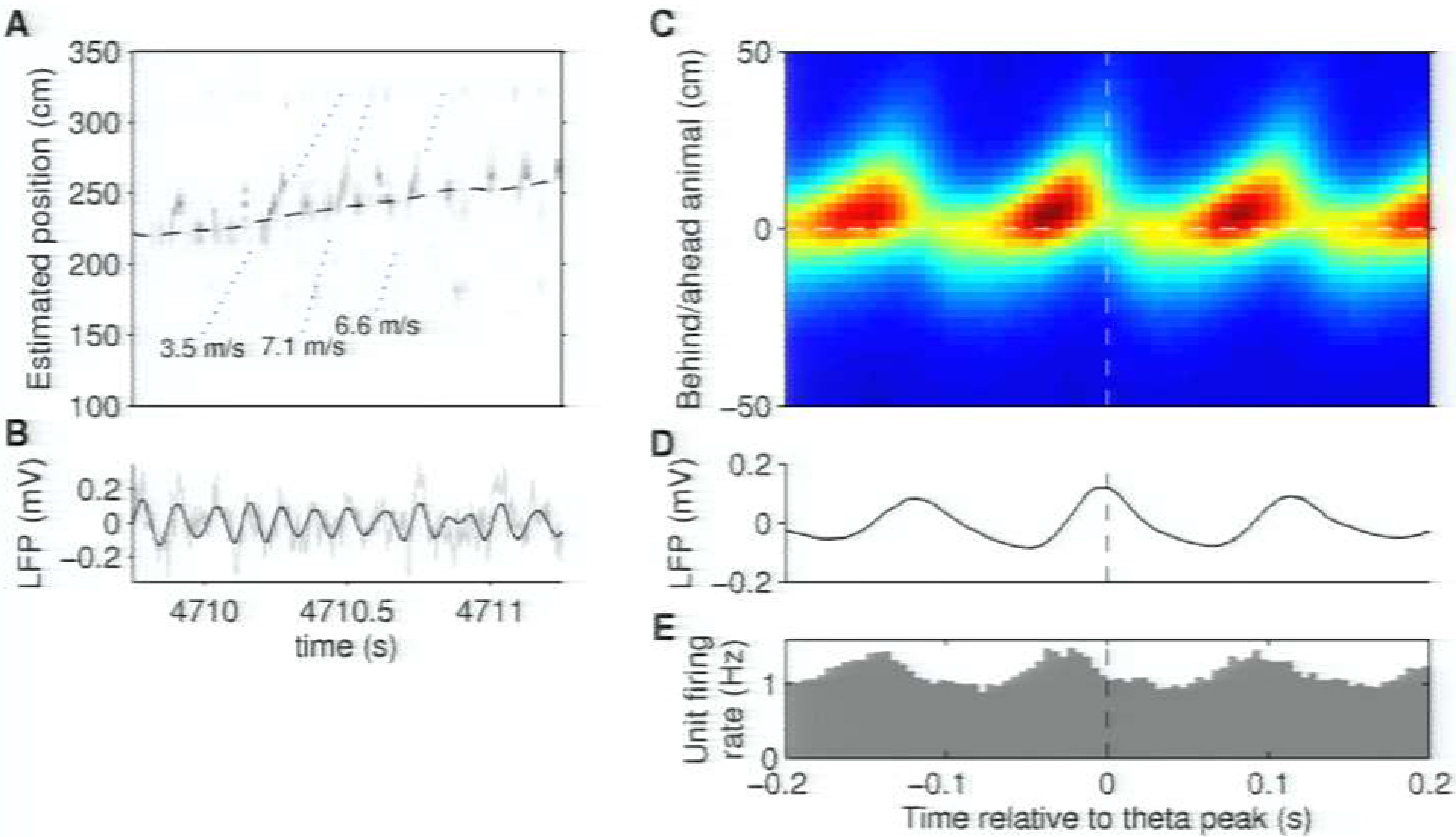
Theta rhythm

Sharp wave/ripples



Buzsaki 1989

Hippocampal spatial representations are encoded as sequences during behavior



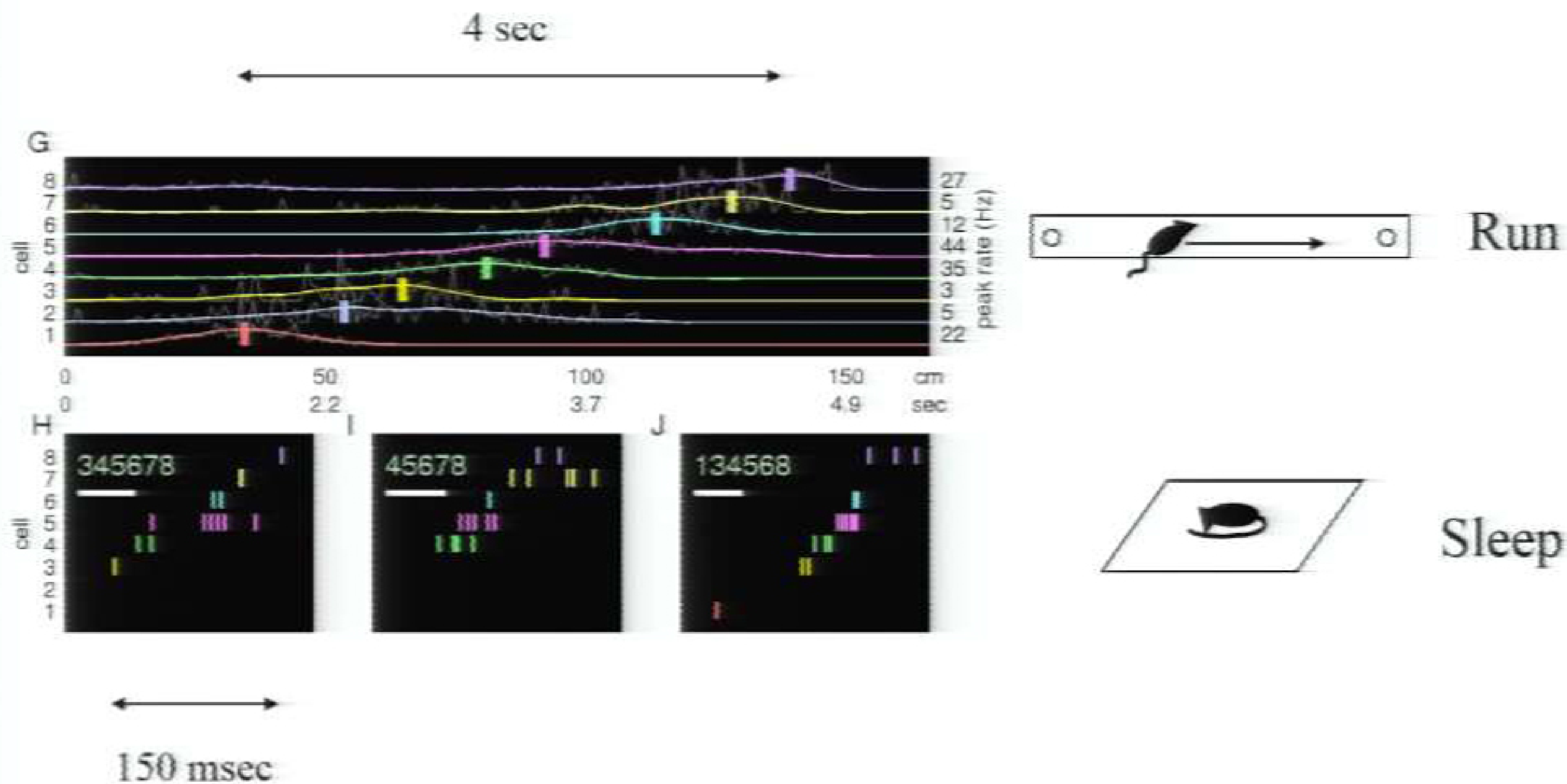
Davidson, Kloosterman, and Wilson, *Neuron*, 2009

Role of Sleep in Memory

- Sleep allows examination of memory independent of behavior.
- The formation of lasting memories may involve the communication of information between brain areas during sleep.
- Broadly identify two stages of non-REM sleep –(NREM) and rapid eye movement sleep (REM).

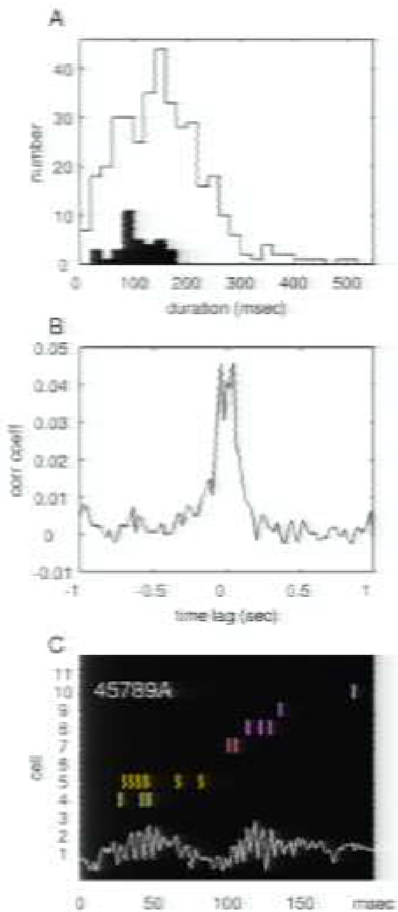
Experimental design

Compressed Run sequences are expressed in hippocampus during nREM sleep



Lee and Wilson, *Neuron*, 2002

Sequences are re-expressed during CA1 ripple events



Duration of low probability sequences

Correlation of low probability sequences and ripples

Example of a low probability sequence and a ripple event

Are there signatures of memory reactivation in the neocortex during hippocampal reactivation

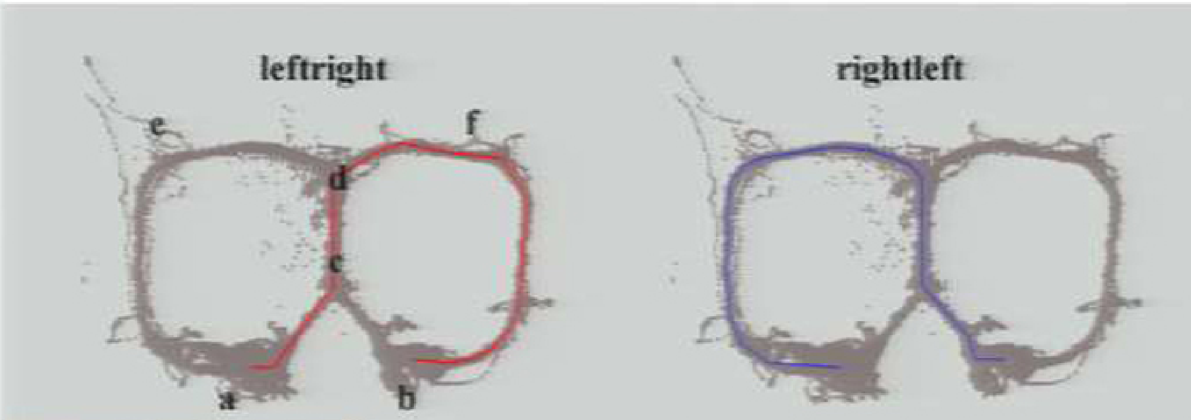
- Simultaneously record in the hippocampus and primary and secondary visual cortex during spatial behavior.
- Look for reactivation in both structures during sleep.

A

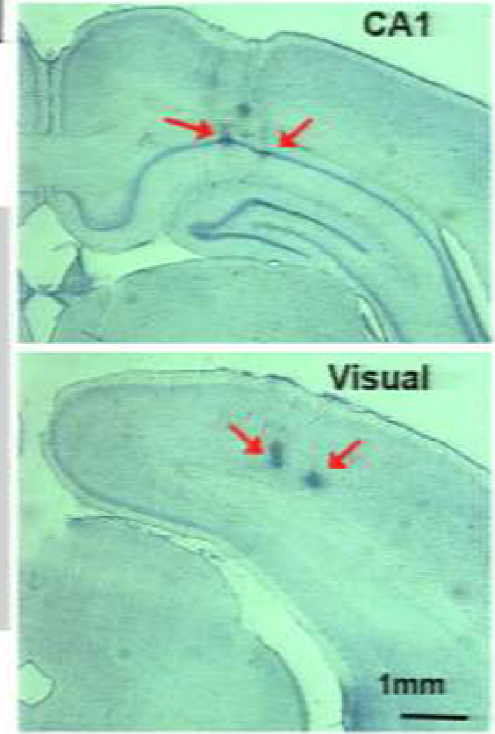
Experimental Design:

PRE (1-2hrs) — RUN (20-40mins) — POST (1-2hrs)

B

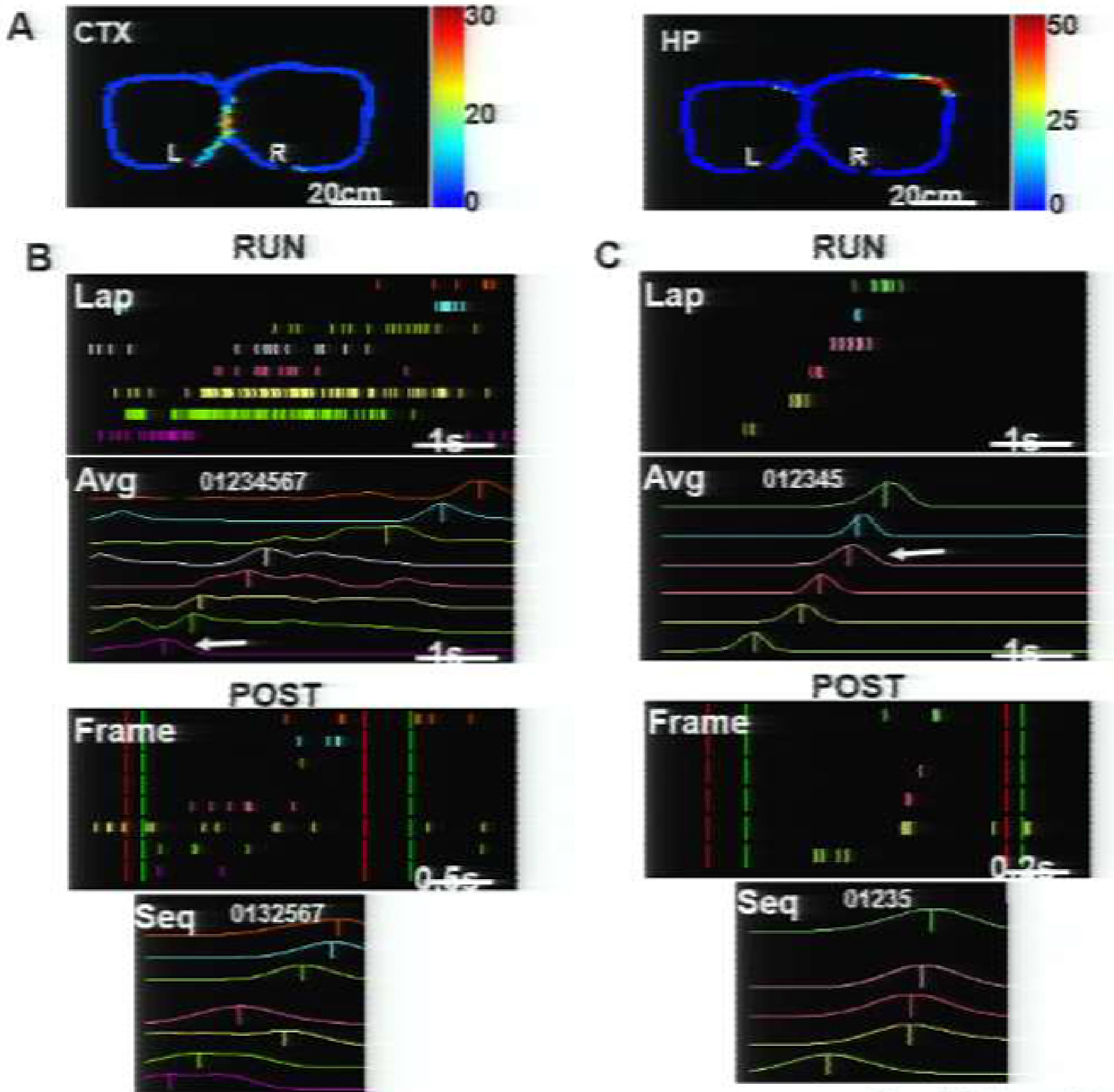


C



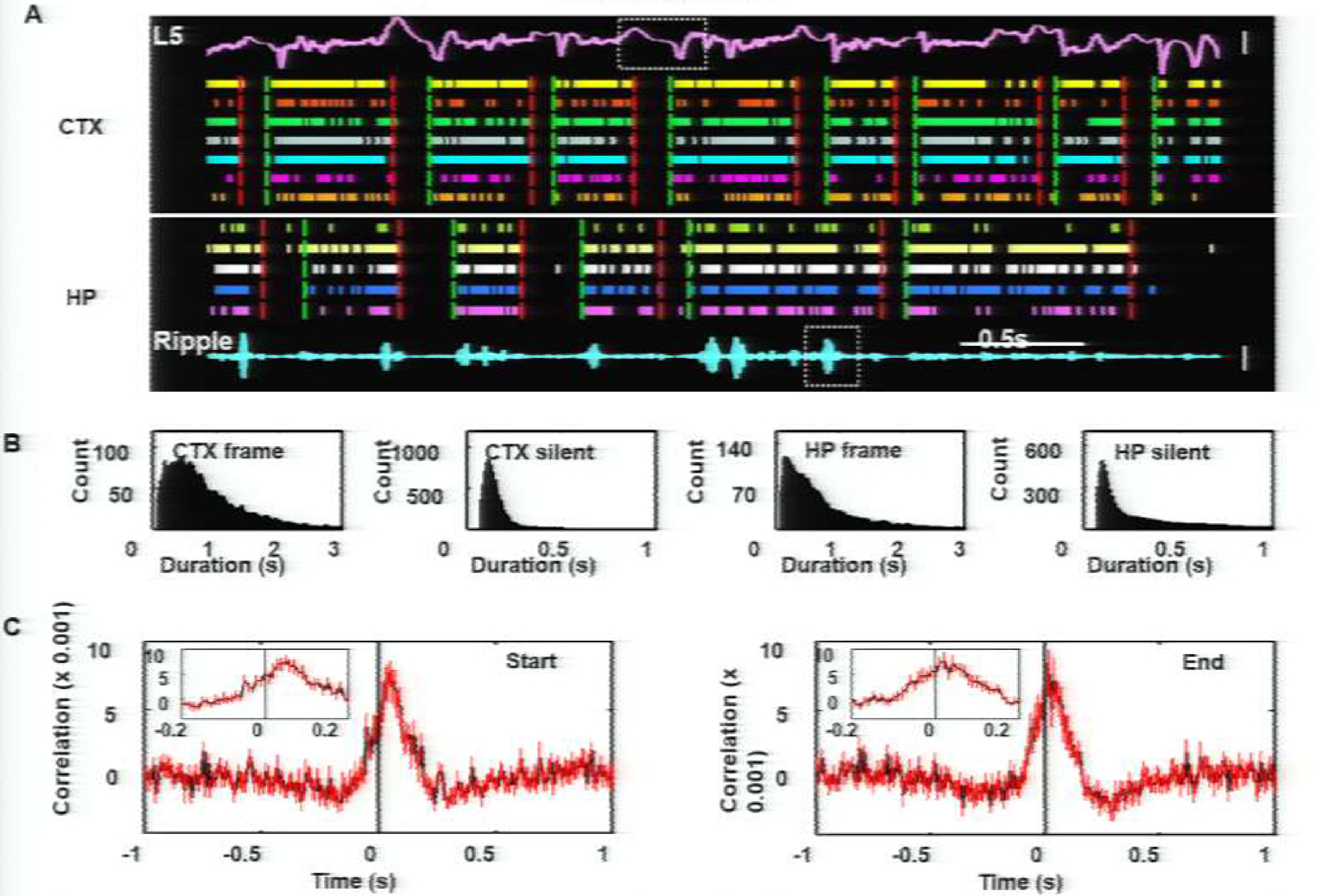
1. Intra-maze local cues, no prominent distal cues
2. Well trained animals: alternation task
3. Recording sites: visual cortex (Occ1, Occ2) and CA1
4. Sleep states (SWS, REM, Wake, Int) classified using EMG and hippocampal EEG

Sequence memory reactivation in hippocampus and visual cortex

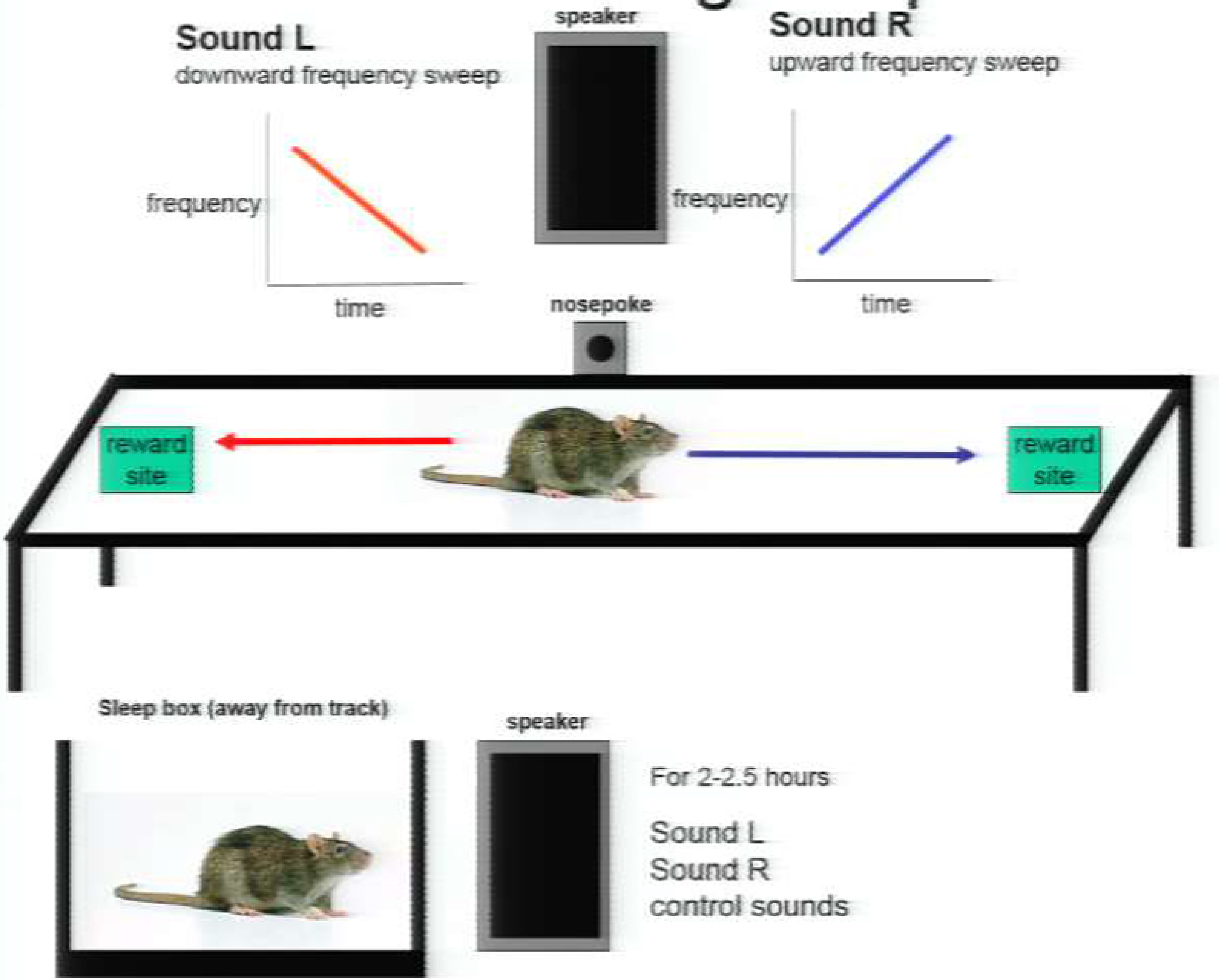


Ji and Wilson, *Nature Neuroscience*, 2007

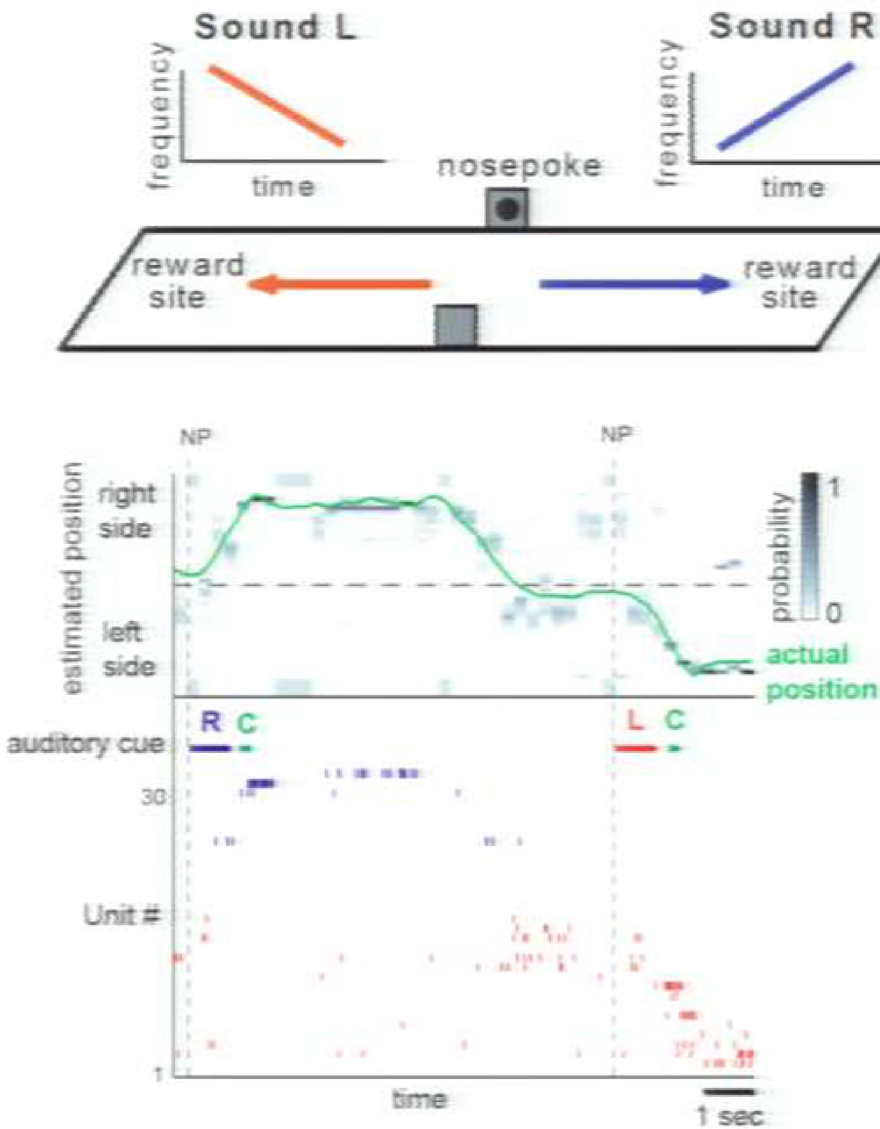
Reactivation occurs during activity frames correlated with the slow oscillation

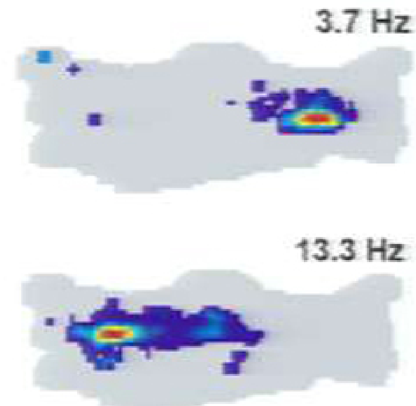


Can we influence memory reactivation during sleep?

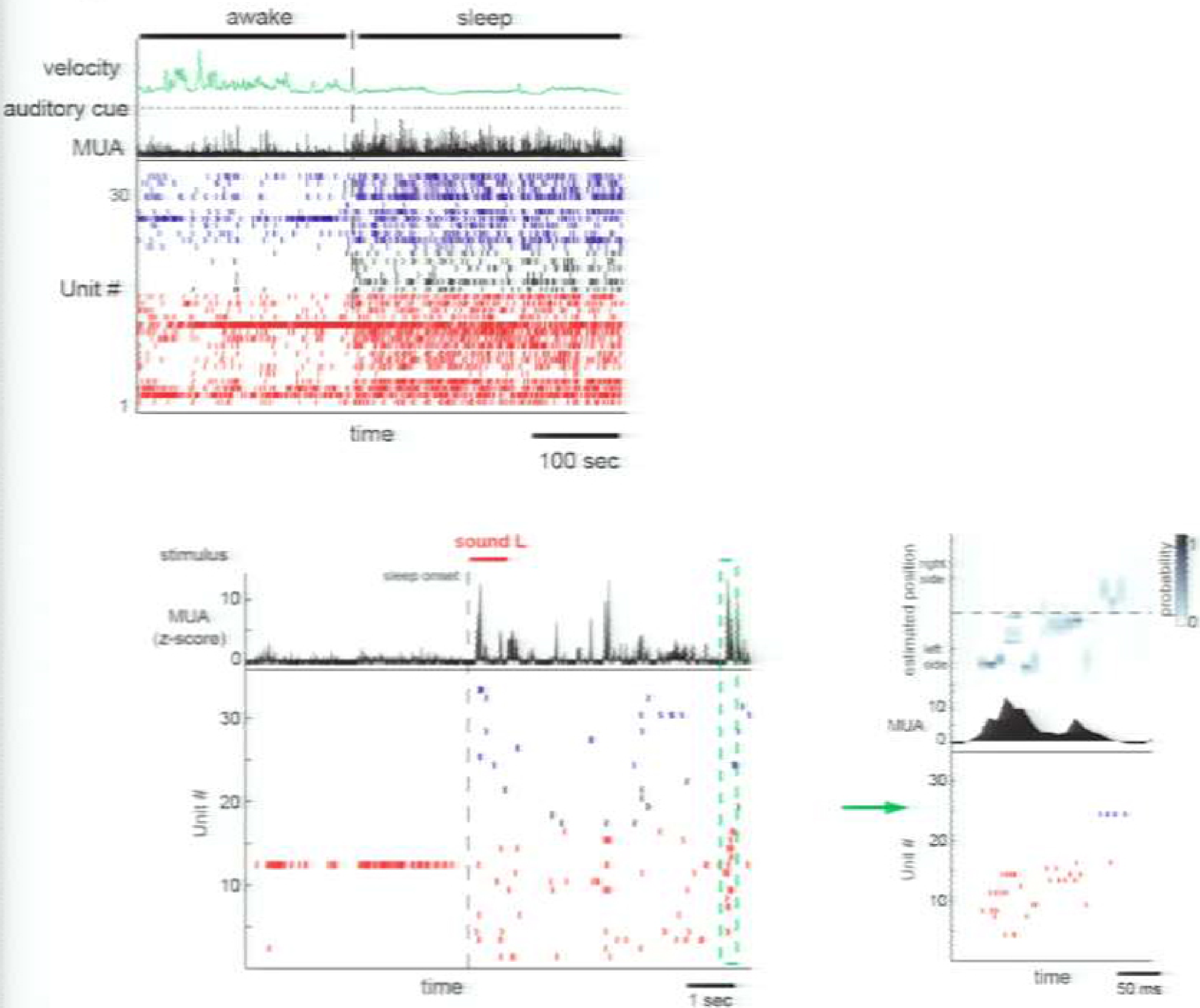


Behavioral task design

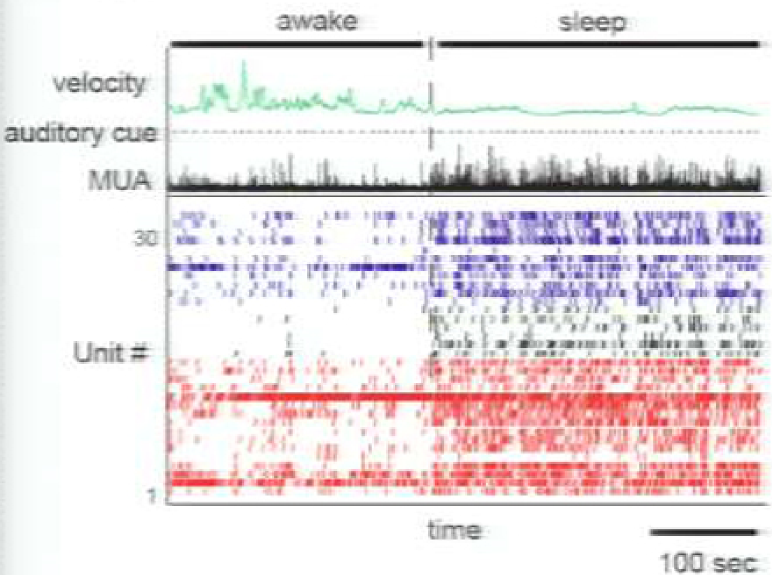




B


awake
sleep
velocity
auditory cue
MUA
30
Unit #
1
time
100 sec
stimulus
sound L
sleep onset
10
MUA
(z-score)
0
30
20
Unit #
10
time
1 sec
estimated position
right side
left side
probability
1
0
MUA
10
0
30
20
Unit #
10
time
50 ms

B

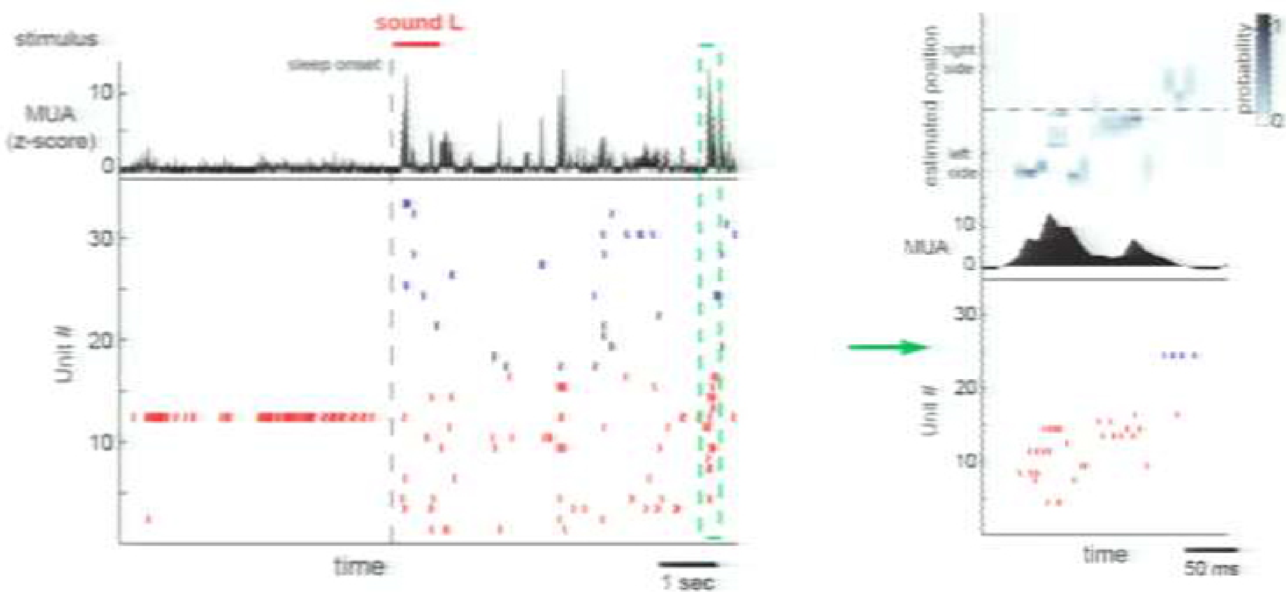


Do task-related sounds bias the content of future replay?

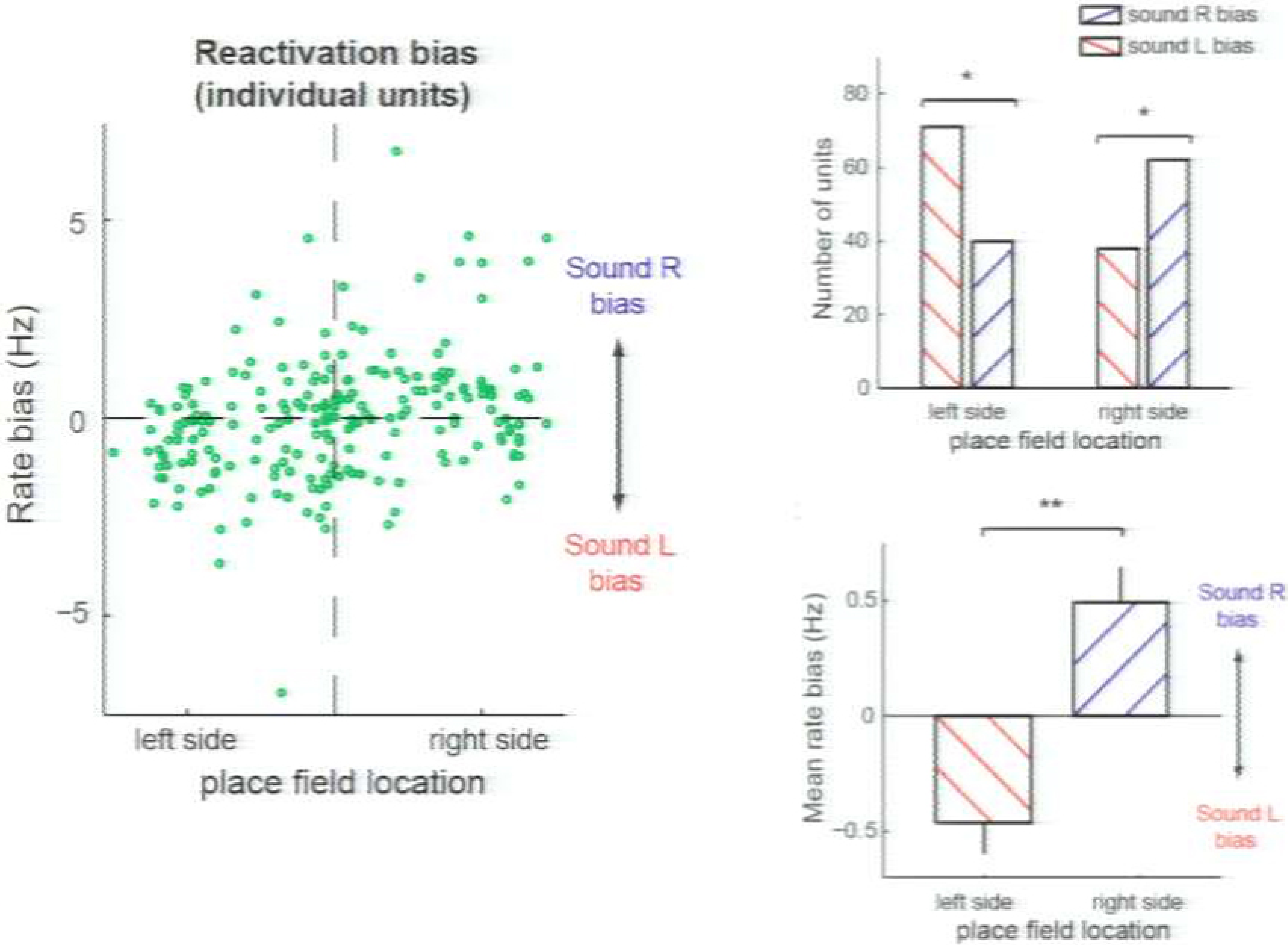
Hypothesis:

Sound R- place cells with **right-sided** place fields are more active during replay

Sound L- place cells with **left-sided** place fields are more active during replay

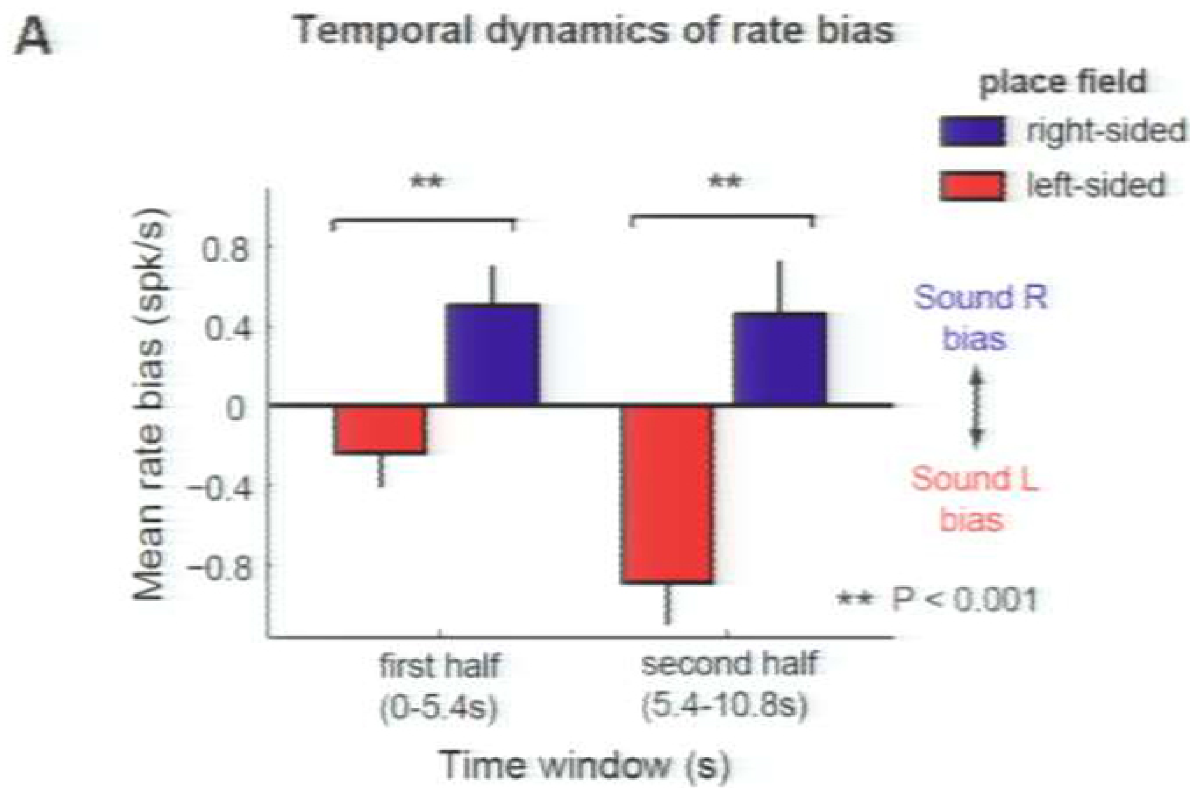


Bias observed in individual place cell responses

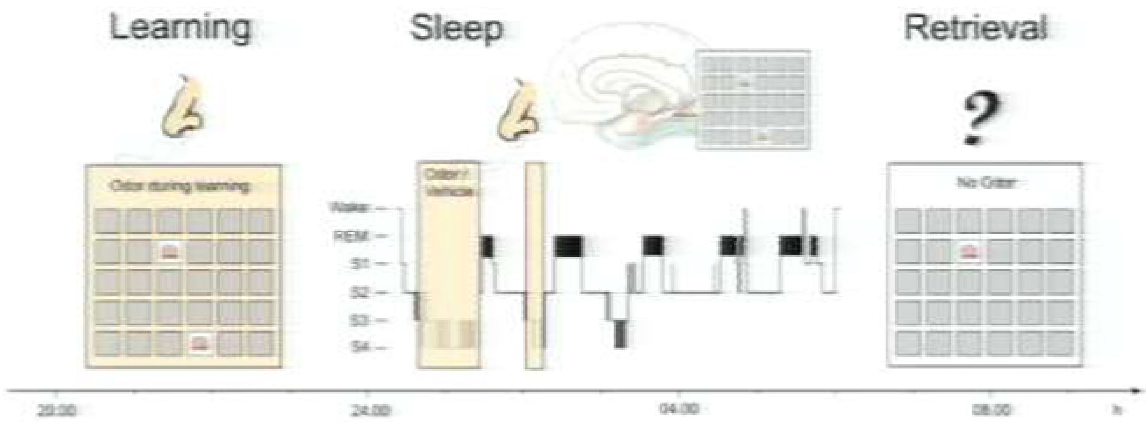


Bendor and Wilson, *Nature Neuroscience*, 2012

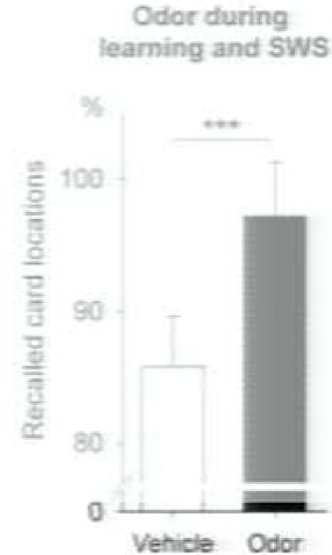
Bias is maintained after initial cueing



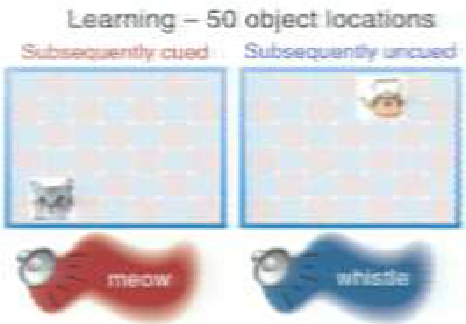
Bendor and Wilson, *Nature Neuroscience*, 2012



Rasch et al. 2007

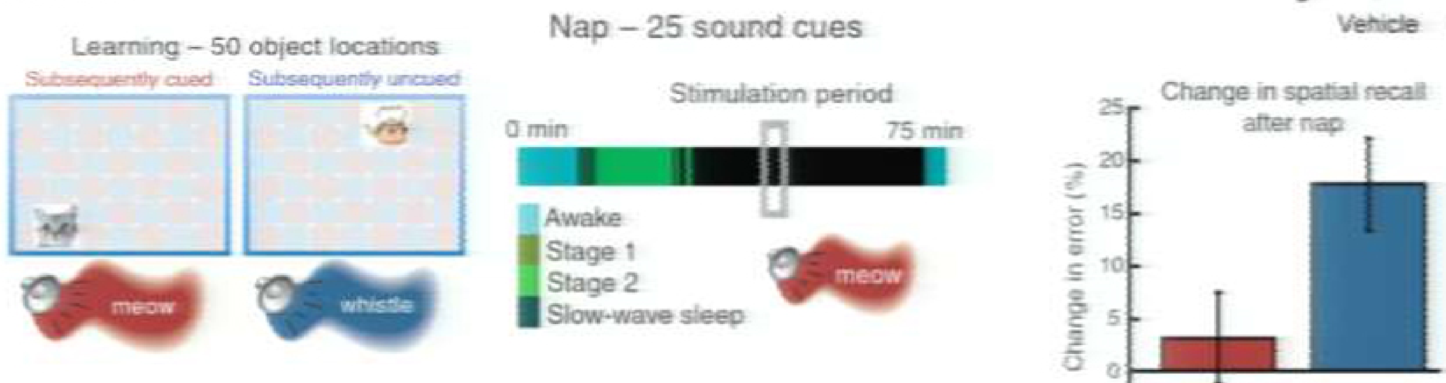


Rasch et al. 2007



Rudoy et al. 2009

Rasch et al. 2007



Rudoy et al. 2009

stimulation ↘

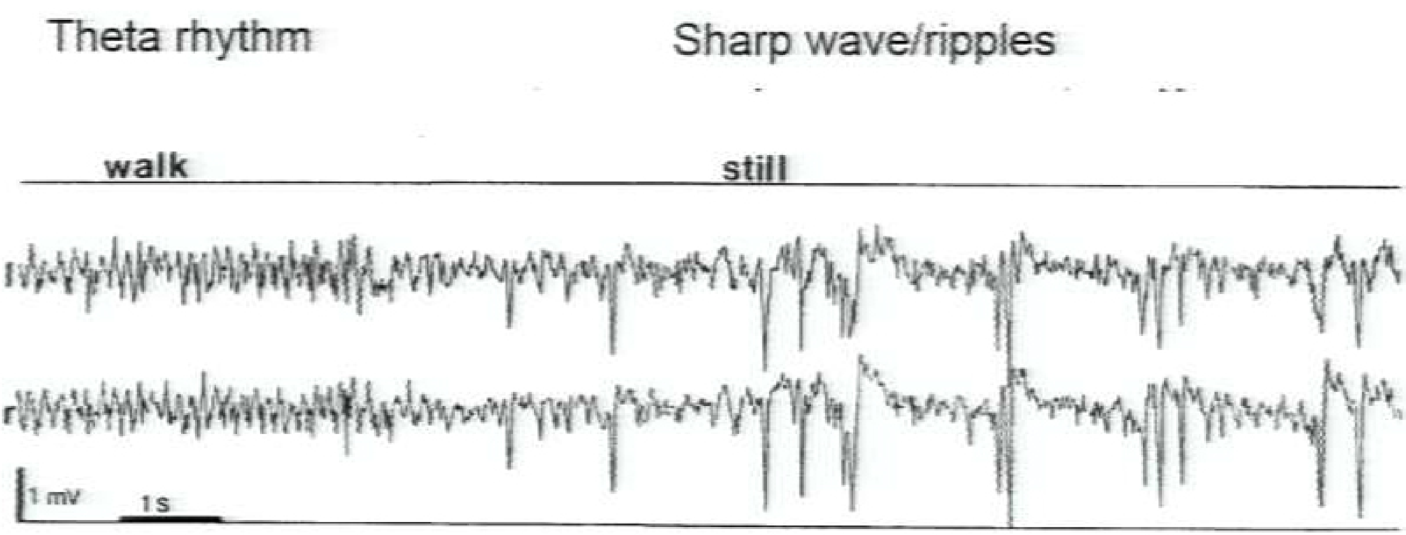
Bias pre-replay brain state

Cortex → Hippocampus

Cortex ← Hippocampus

Bias which memories are transferred

Hippocampus online and offline

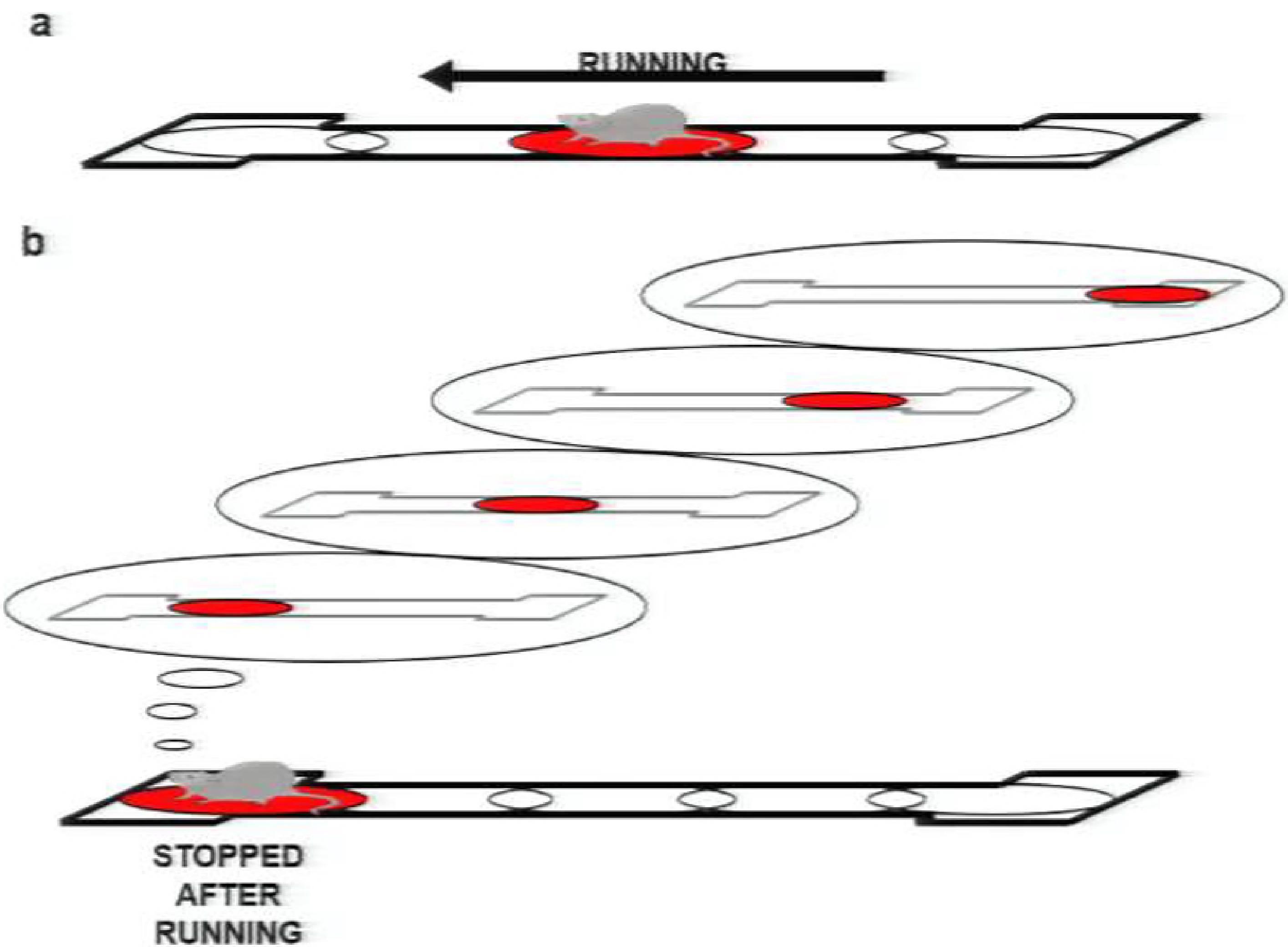


Buzsaki 1989

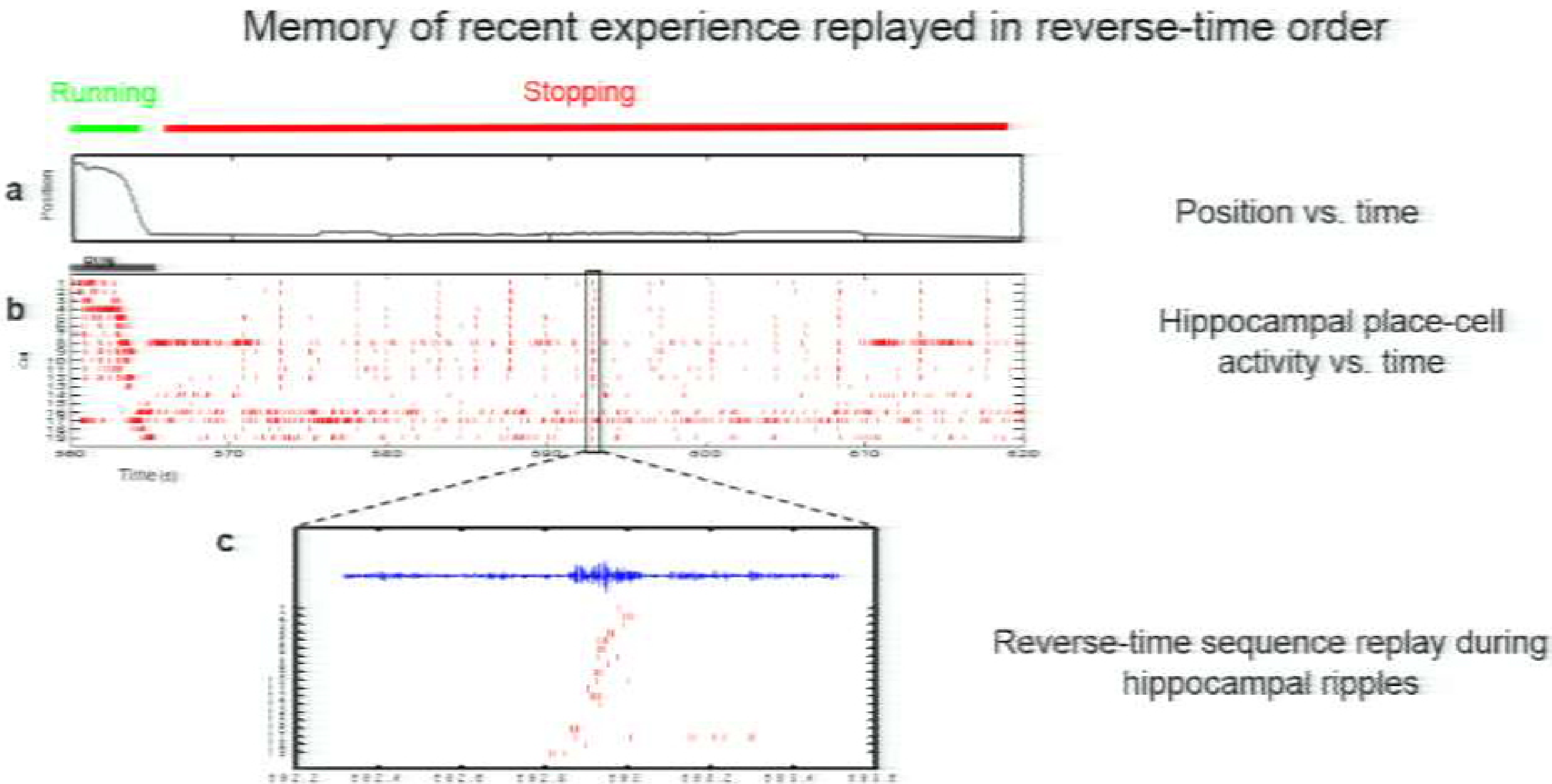
Hippocampal activity during quiet wakefulness

- During awake behavior, there are periods of quiet wakefulness that have EEG that is similar to NREM consisting of brief bursts of activity modulated by high frequency "ripple" oscillations.
- Is there structure to the patterns of multiple single neuron activity during this state?

What do animals think about when they stop and eat after running down a track?



They think back to where they have just been.

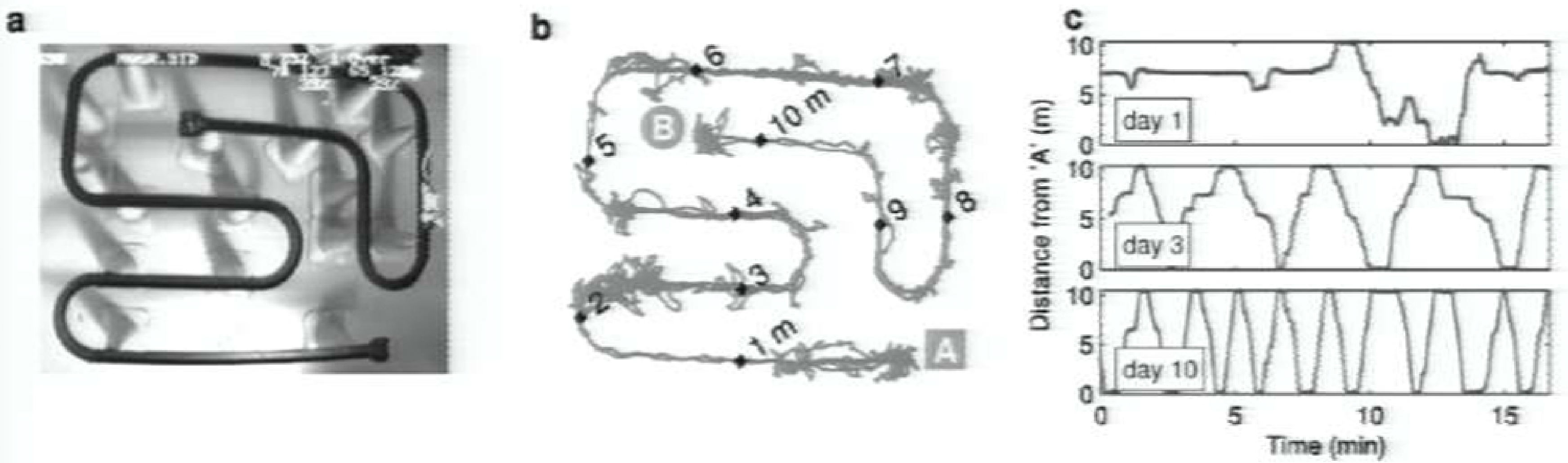


Foster and Wilson, *Nature*, 2006

Questions

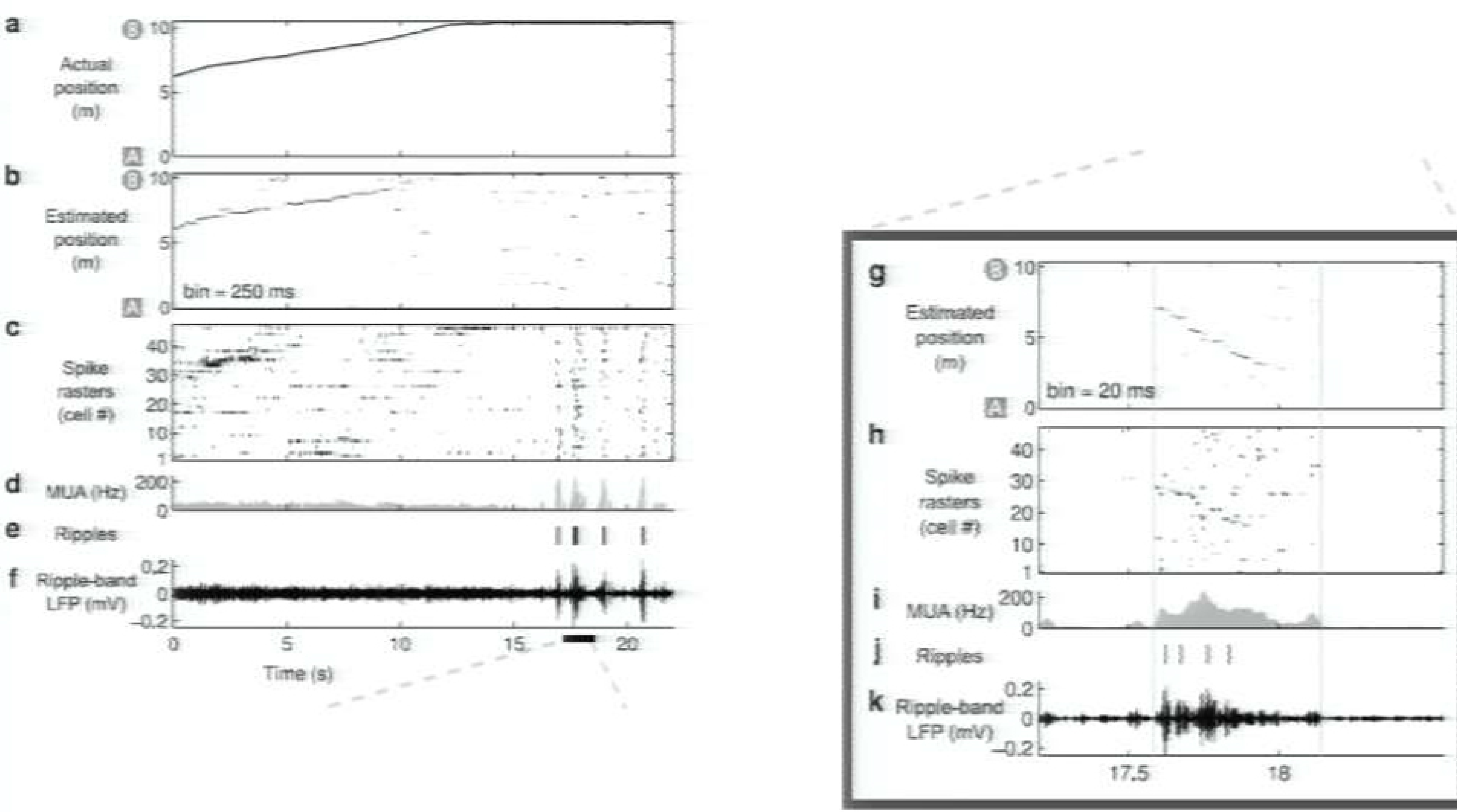
- ✓ Replay in a larger environment?
- ✓ Replay associated with reward sites only?
- ✓ Replay always begins with cells that have place fields close to animal's current location?
- ✓ Replay in forward and reverse directions?

Long behavioral sequences on a 10m track



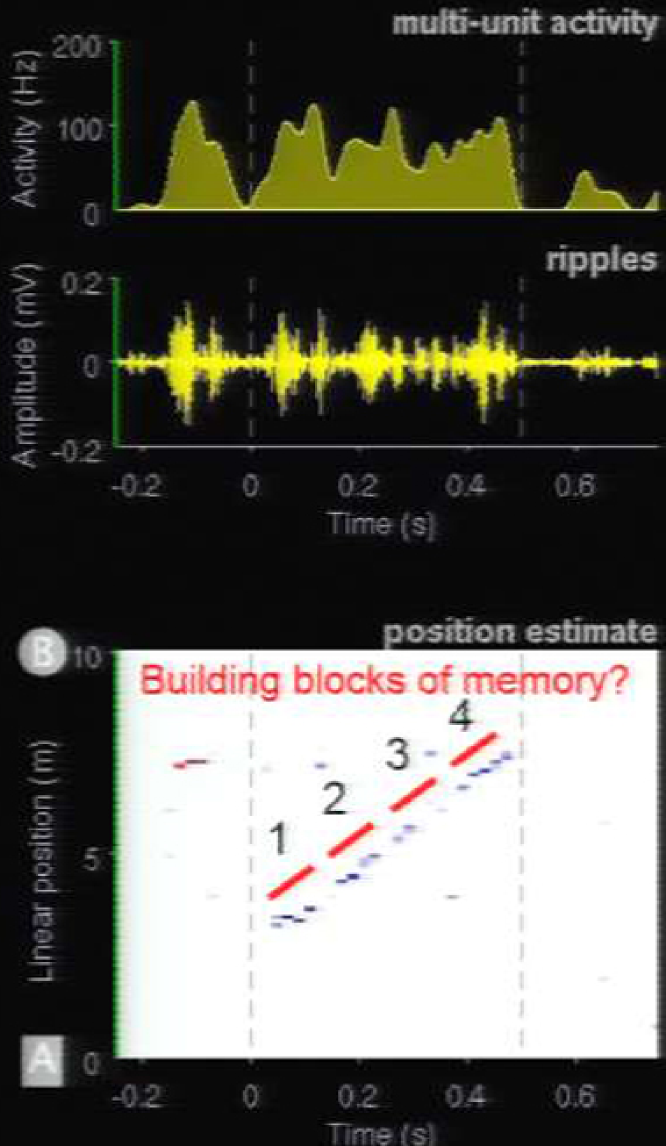
Davidson, Kloosterman and Wilson, *Neuron*, 2009

Reconstruction of extended sequence replay during quiet wakefulness

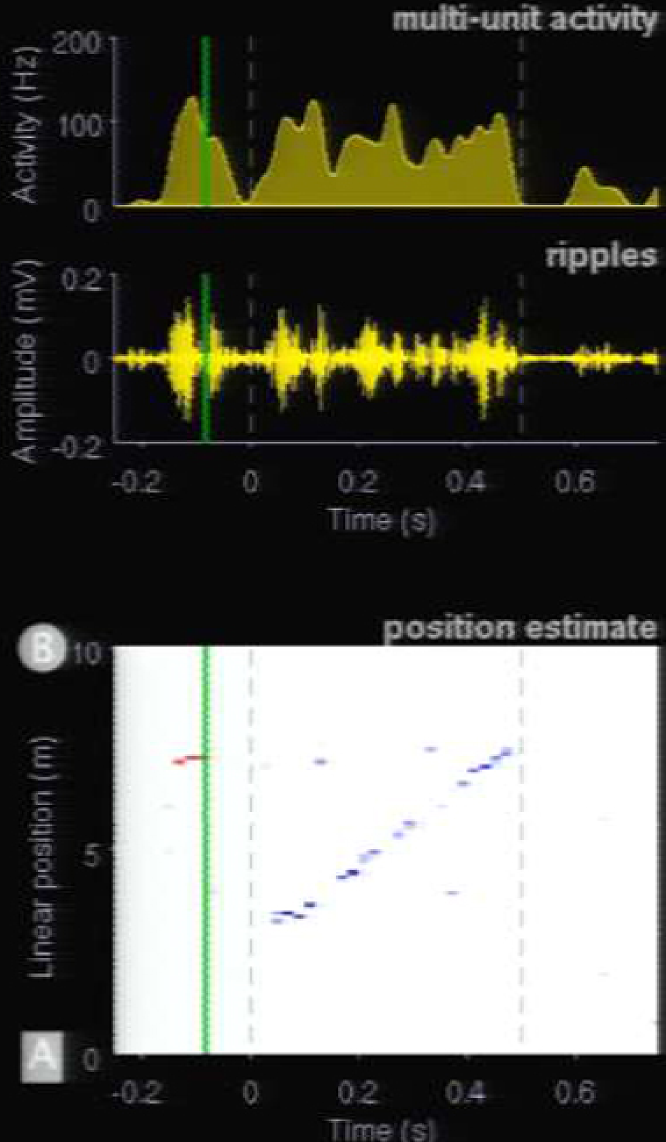


Davidson, Kloosterman and Wilson, *Neuron*, 2009

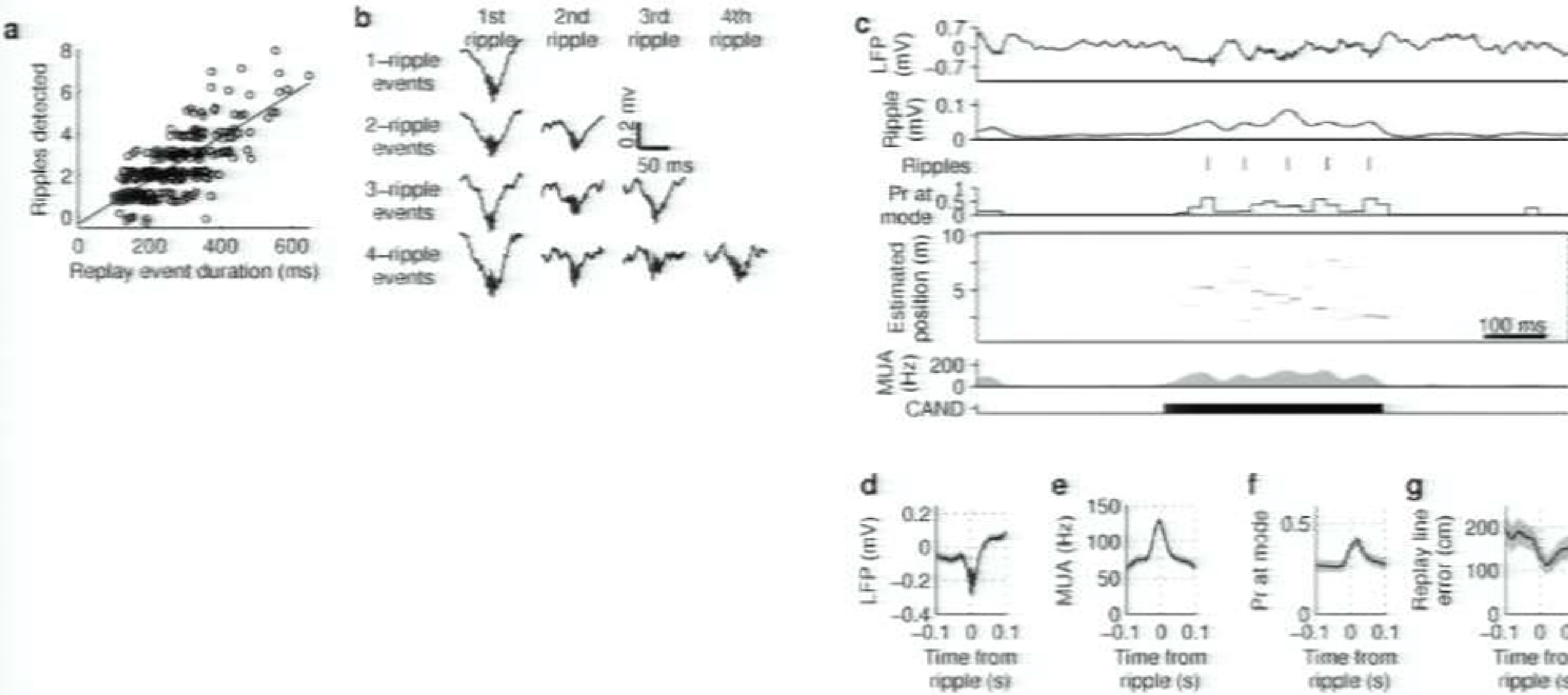
Forward Replay from A to B



Forward Replay from A to B

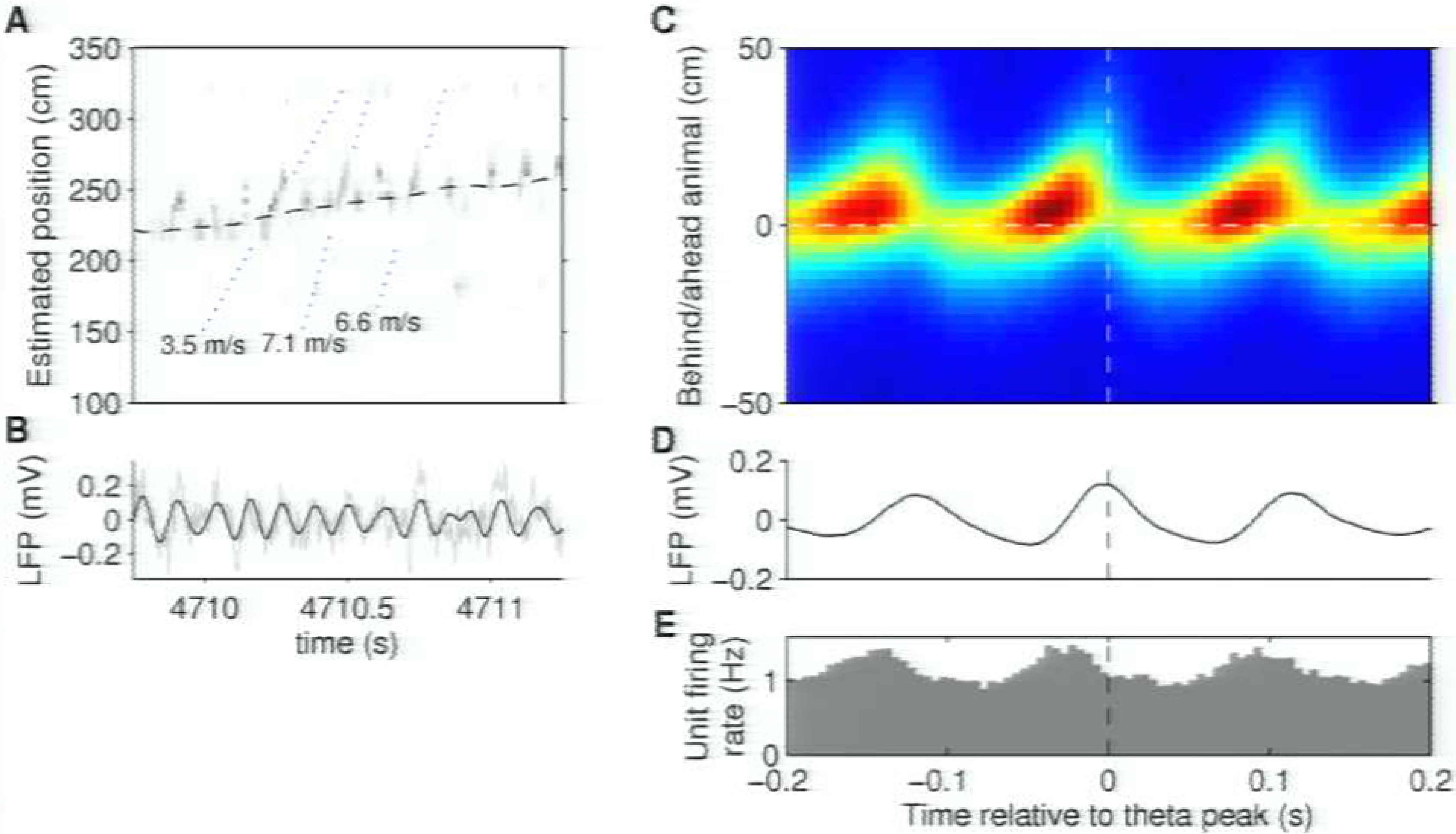


Extended replay spans multiple ripple events



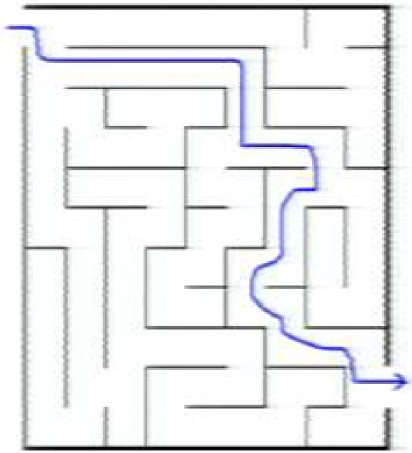
Davidson, Kloosterman and Wilson, *Neuron*, 2009

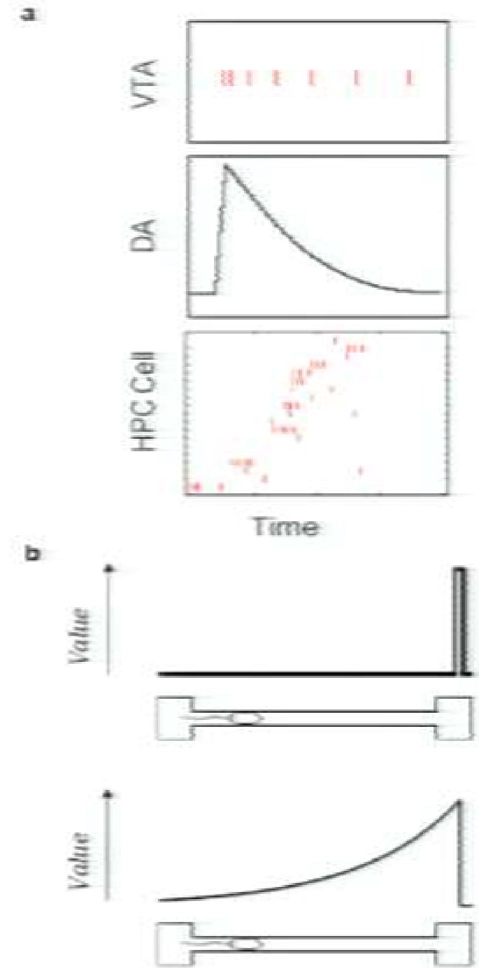
Single ripple sequences are at same scale as theta sequences



Davidson, Kloosterman, and Wilson, *Neuron*, 2009

Learning sequences of actions



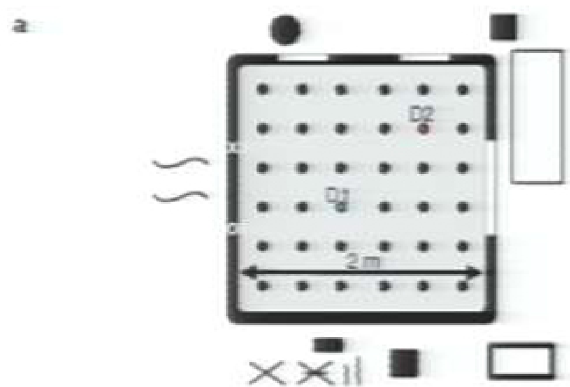


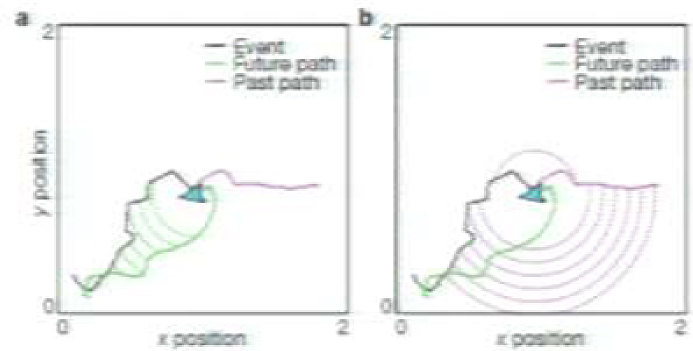
Temporal credit assignment

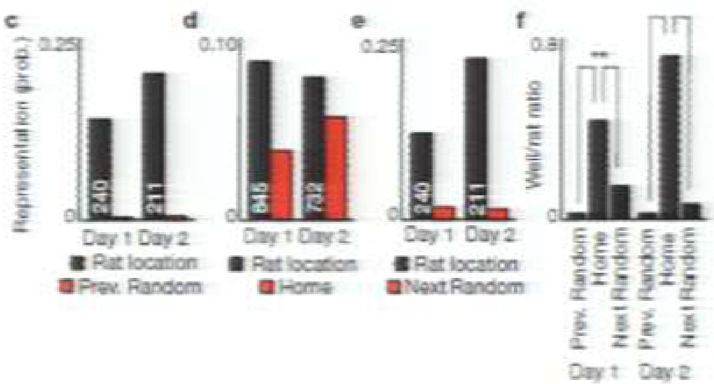
Dopamine unit activity could differentially weight the content of hippocampal sequences, propagating value information from the rewarded location backwards along the incoming trajectory.

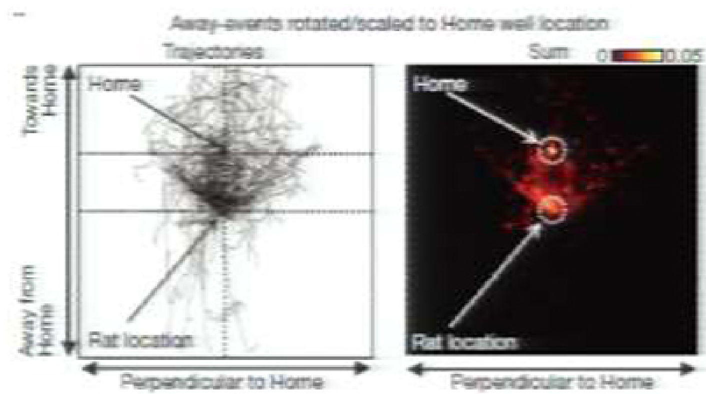
Foster and Wilson, Nature 440: 680-683, 2006

Hippocampal place-cell sequences depict future paths to remembered goals
Brad E. Pfeiffer & David J. Foster
Nature, 2013

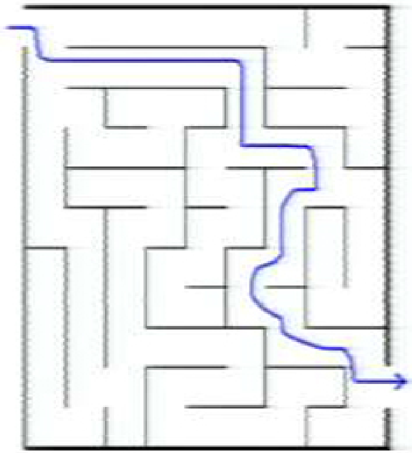


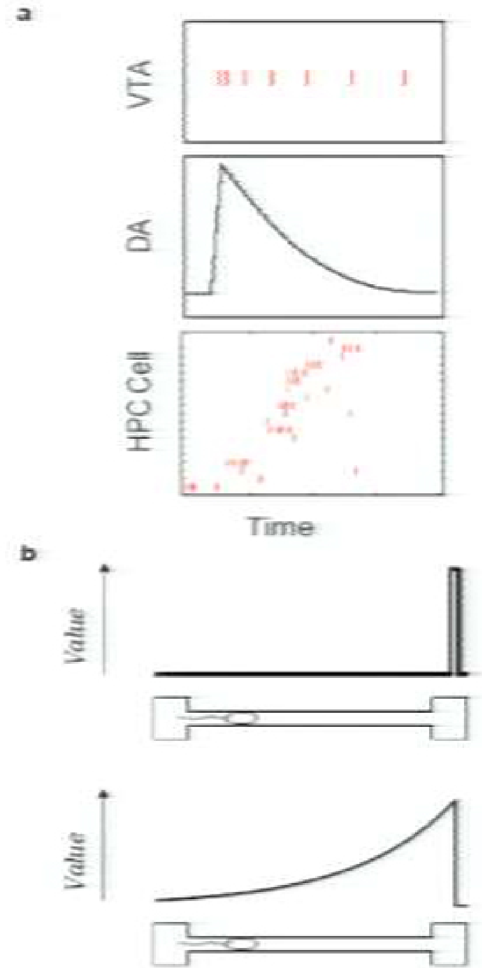






Learning sequences of actions





Temporal credit assignment

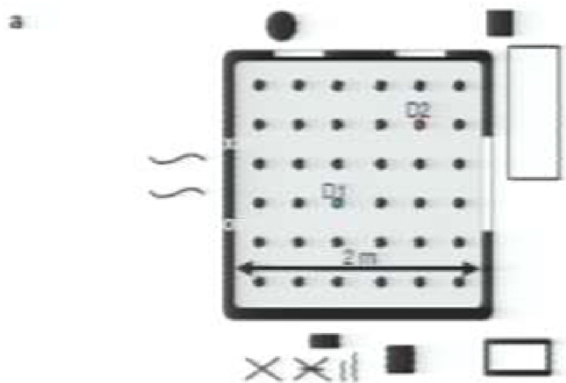
Dopamine unit activity could differentially weight the content of hippocampal sequences, propagating value information from the rewarded location backwards along the incoming trajectory.

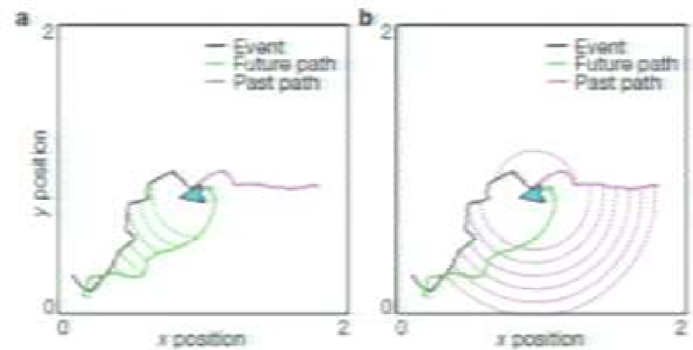
Foster and Wilson, Nature 440: 680-683, 2006

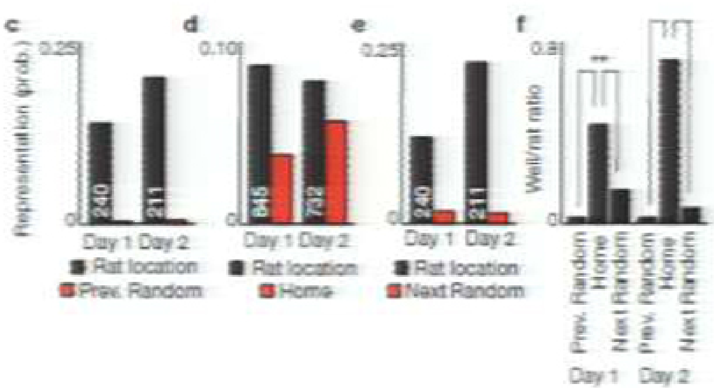
Hippocampal place-cell sequences depict future paths to remembered goals

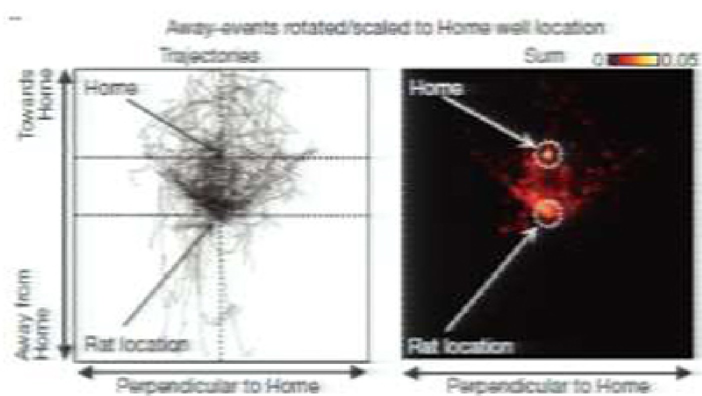
Brad E. Pfeiffer & David J. Foster

Nature, 2013



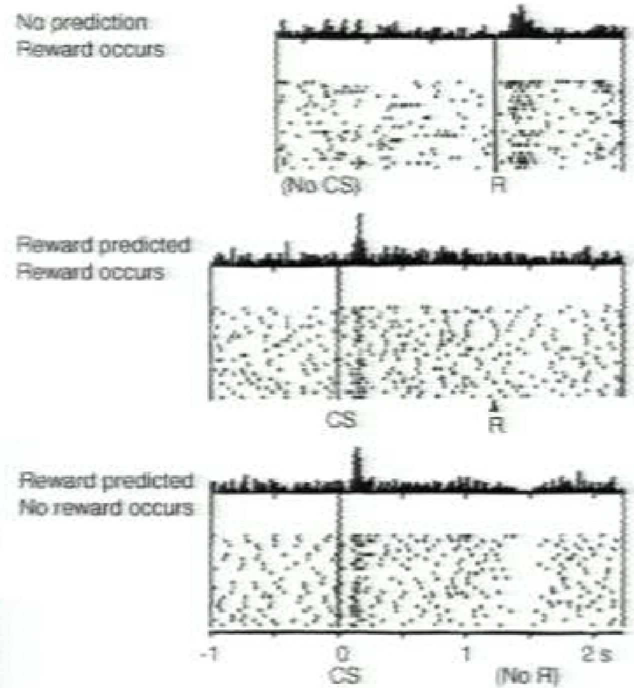






Dopamine cell representations

- unexpected reward
- predictors of reward
- errors in the prediction of reward



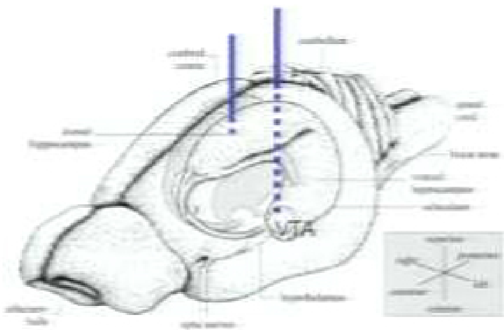
Reward prediction error

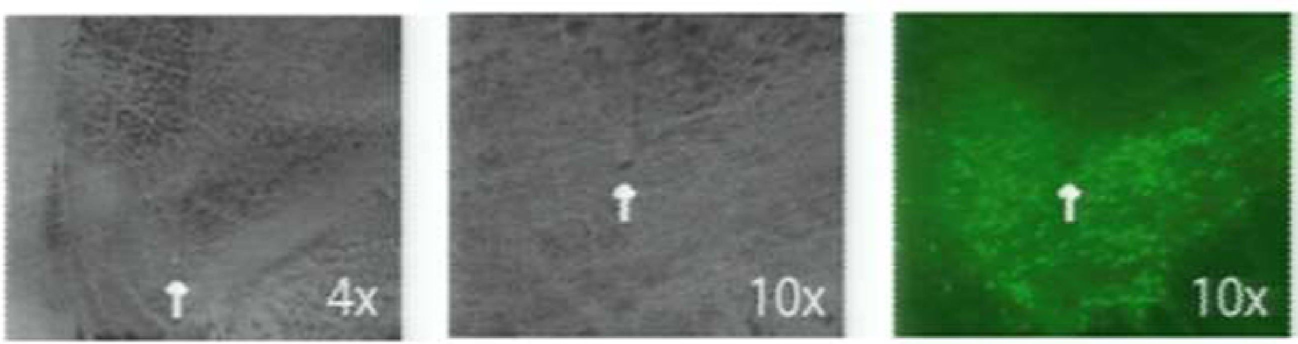
Current reward – Expected reward

An error signal
to teach target
brain regions

Schultz, Dayan, Montague. Science 1997

VTA Hippocampus co-recordings

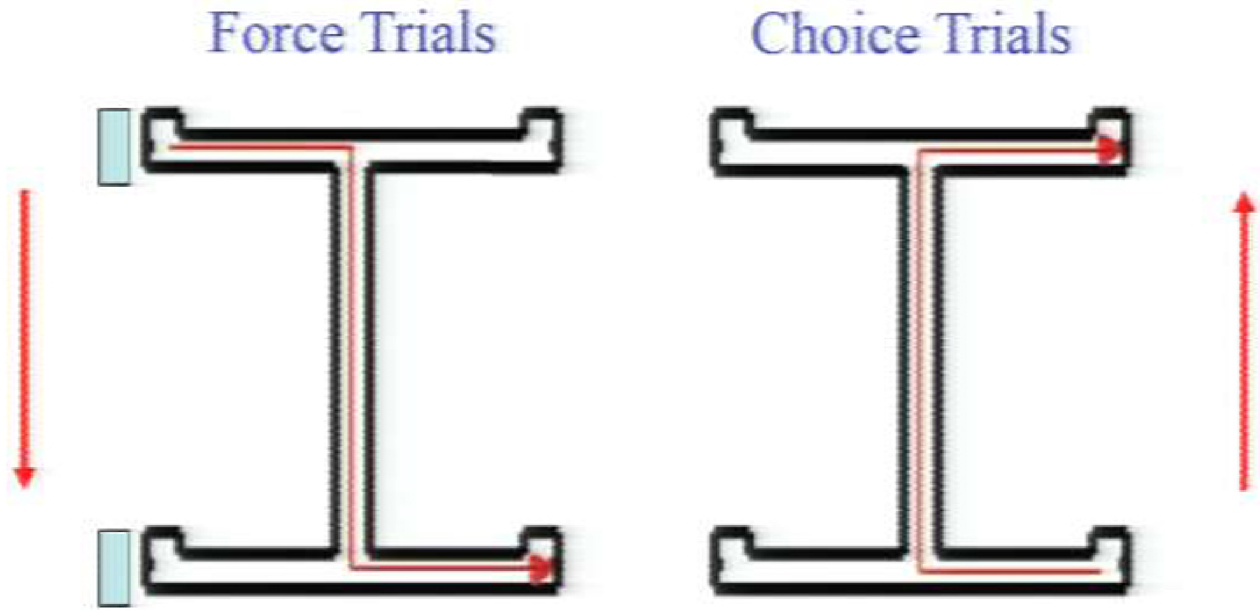




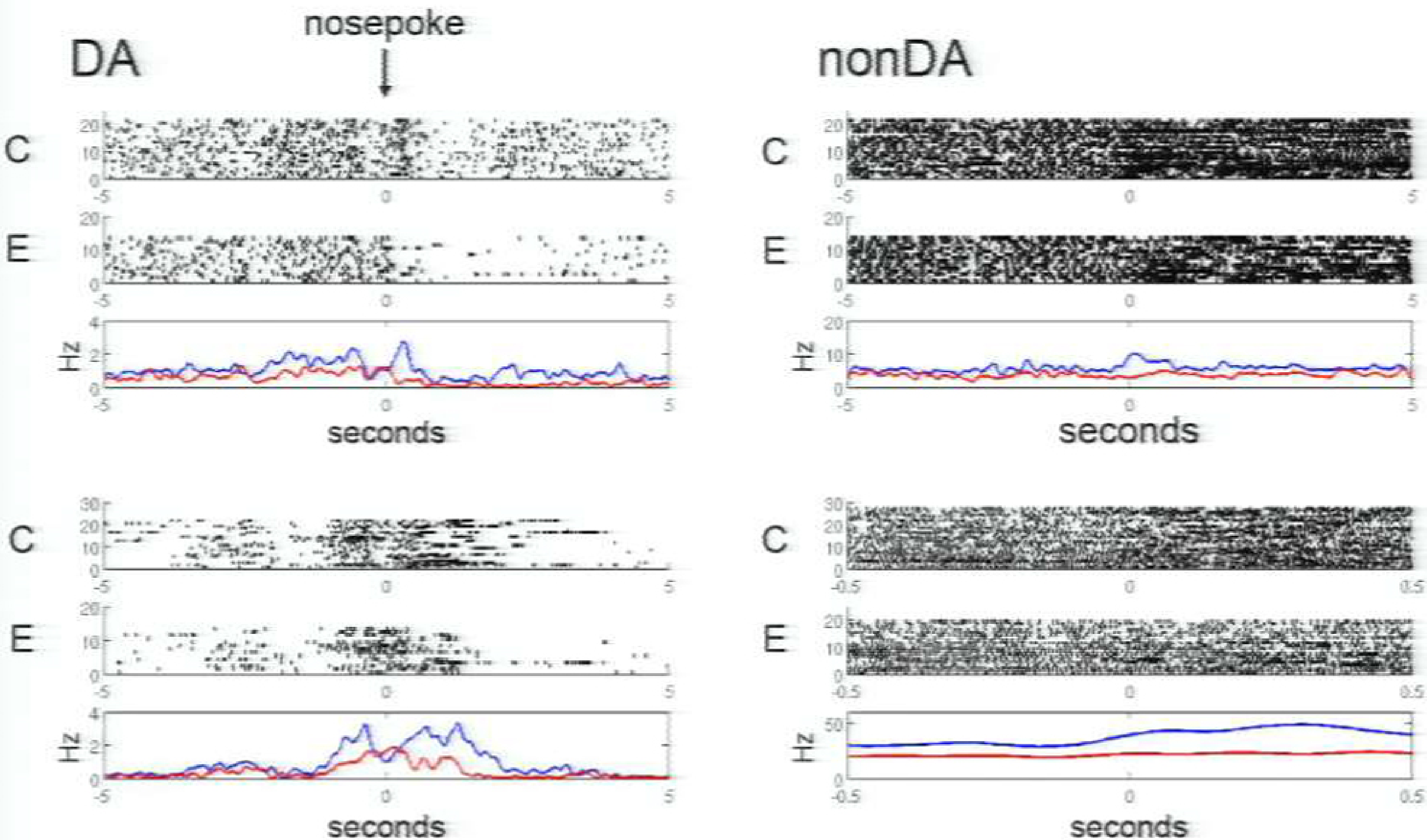
Light microscopy

Anti-TH Ab

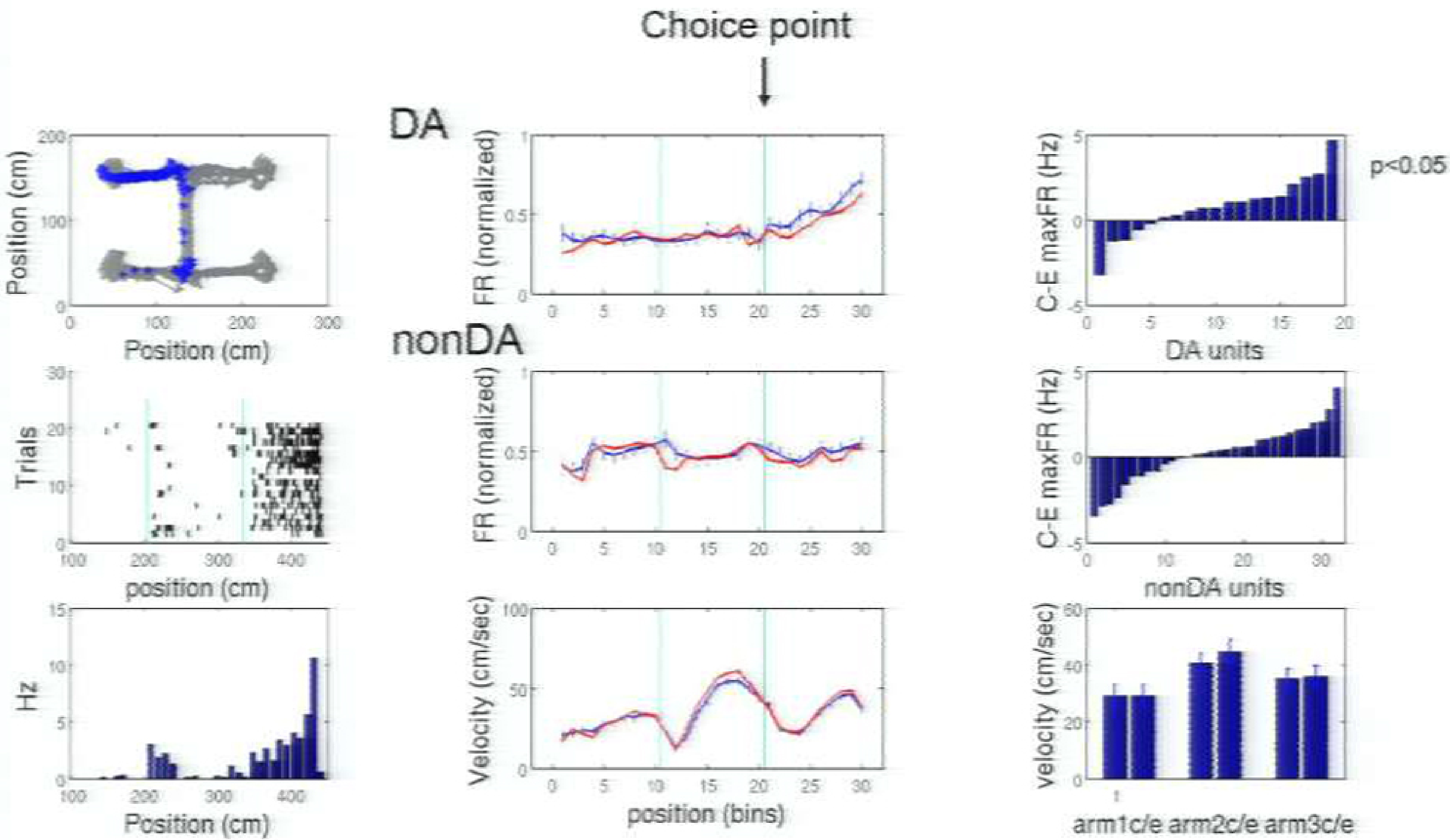
Spatial working memory task



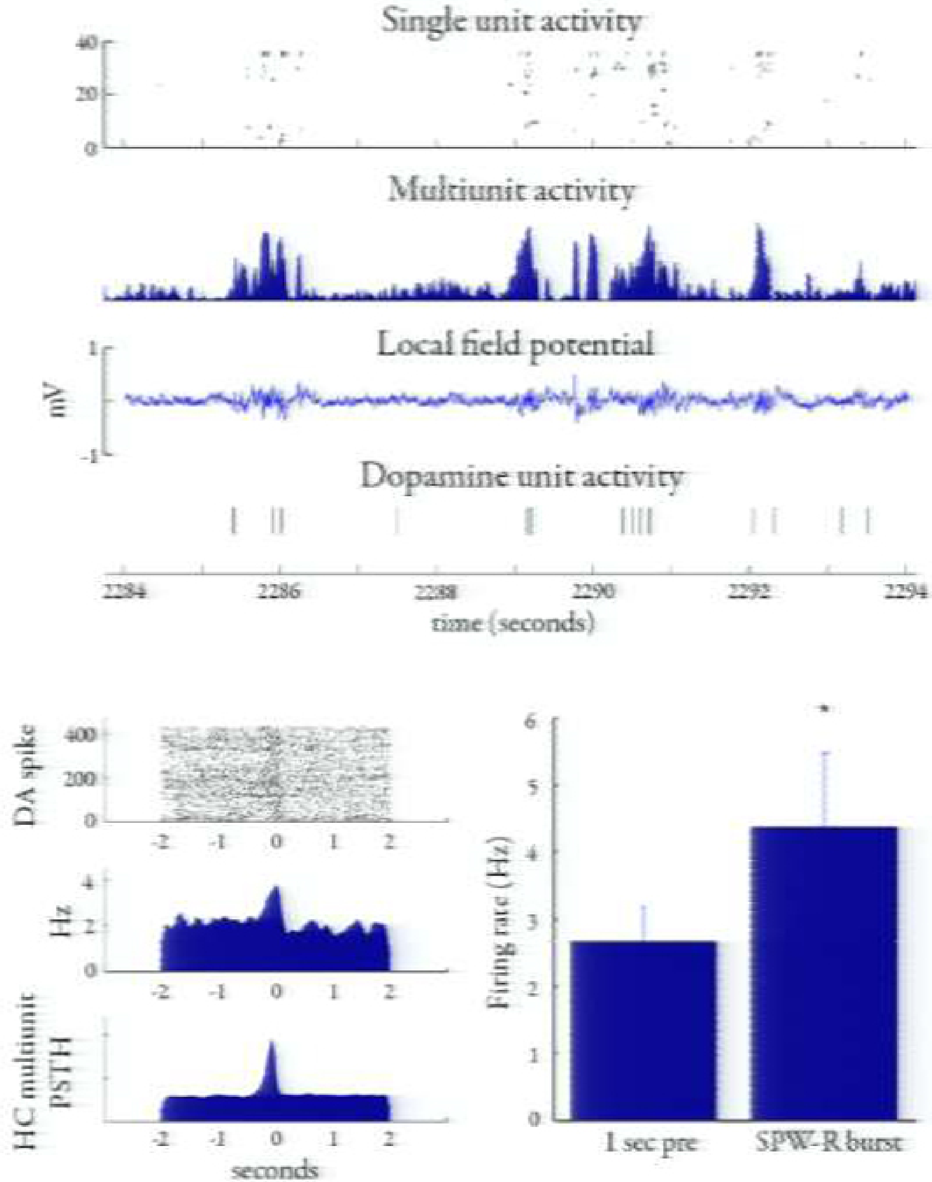
Task contingency associated VTA unit activity



VTA unit activity during RUN

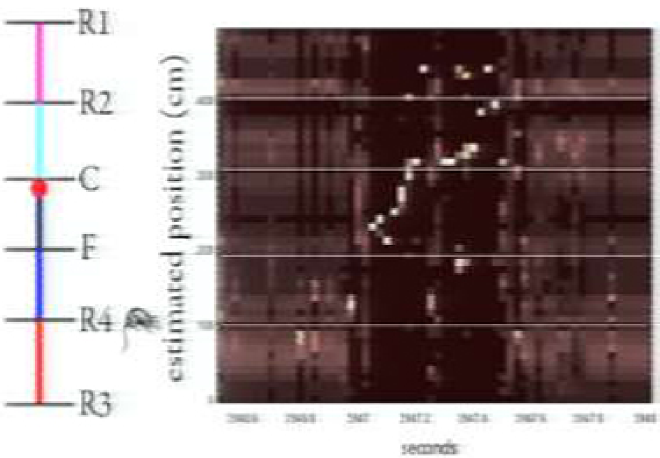


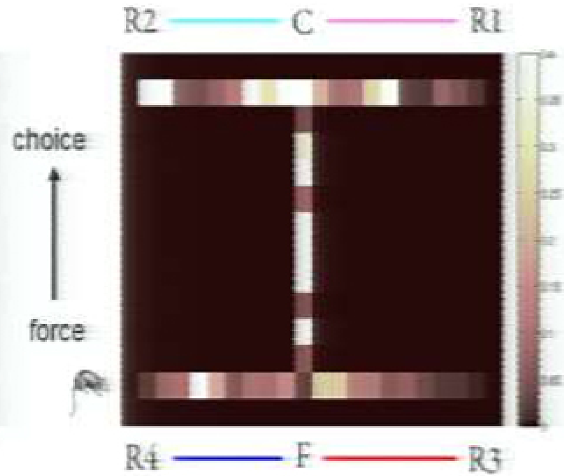
HC ripple bursts modulate DA unit activity

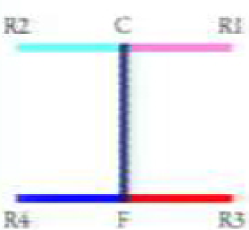


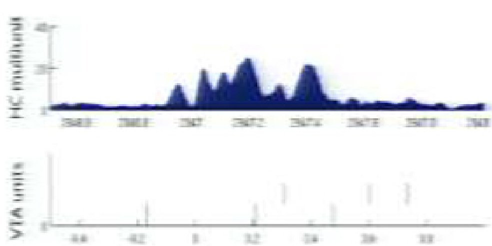
Gomperts, unpublished data

Decoding hippocampal SPW-R associated multiunit bursts with spatial sequence reactivation

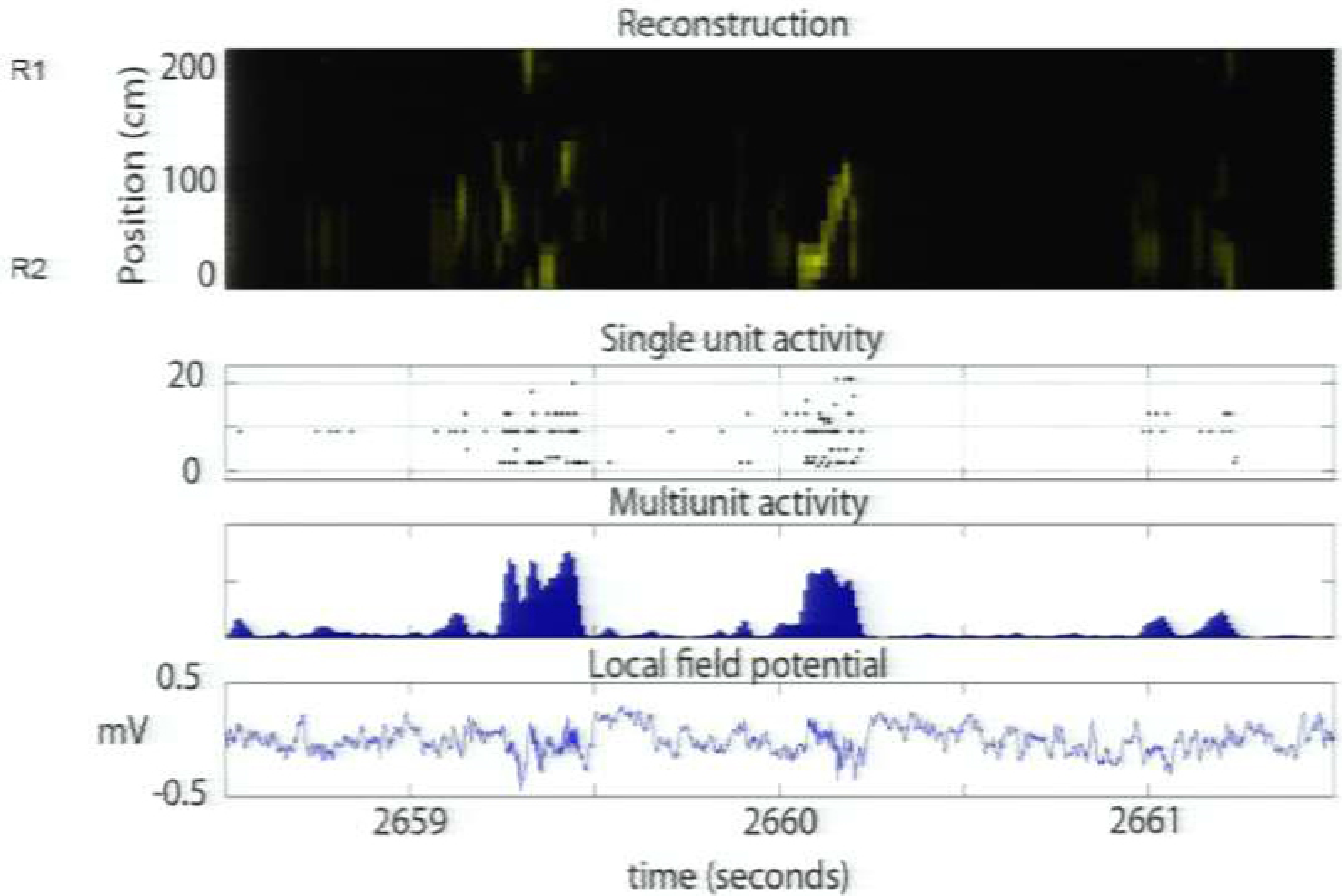






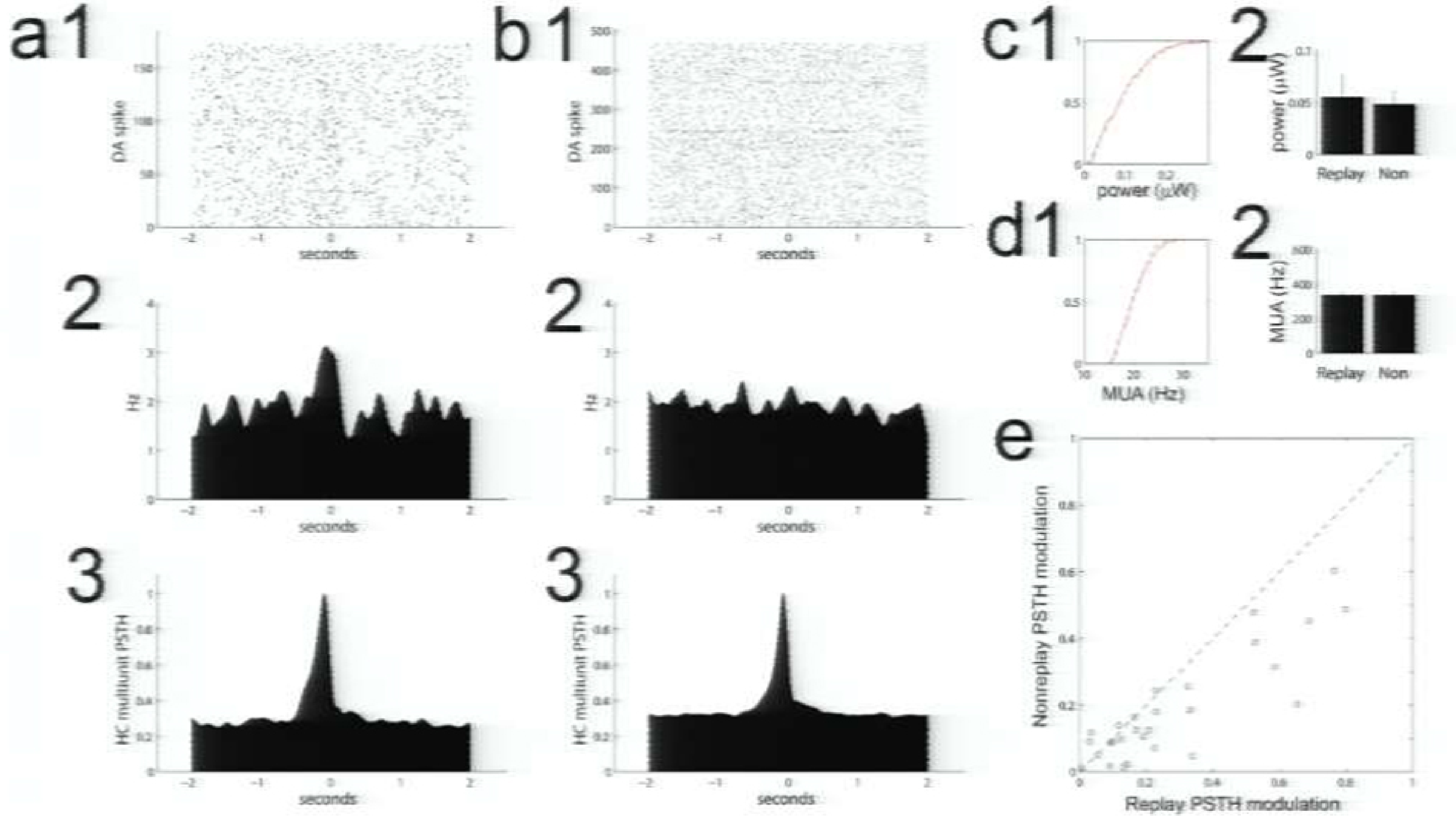


Replay and nonreplay SPW-R events

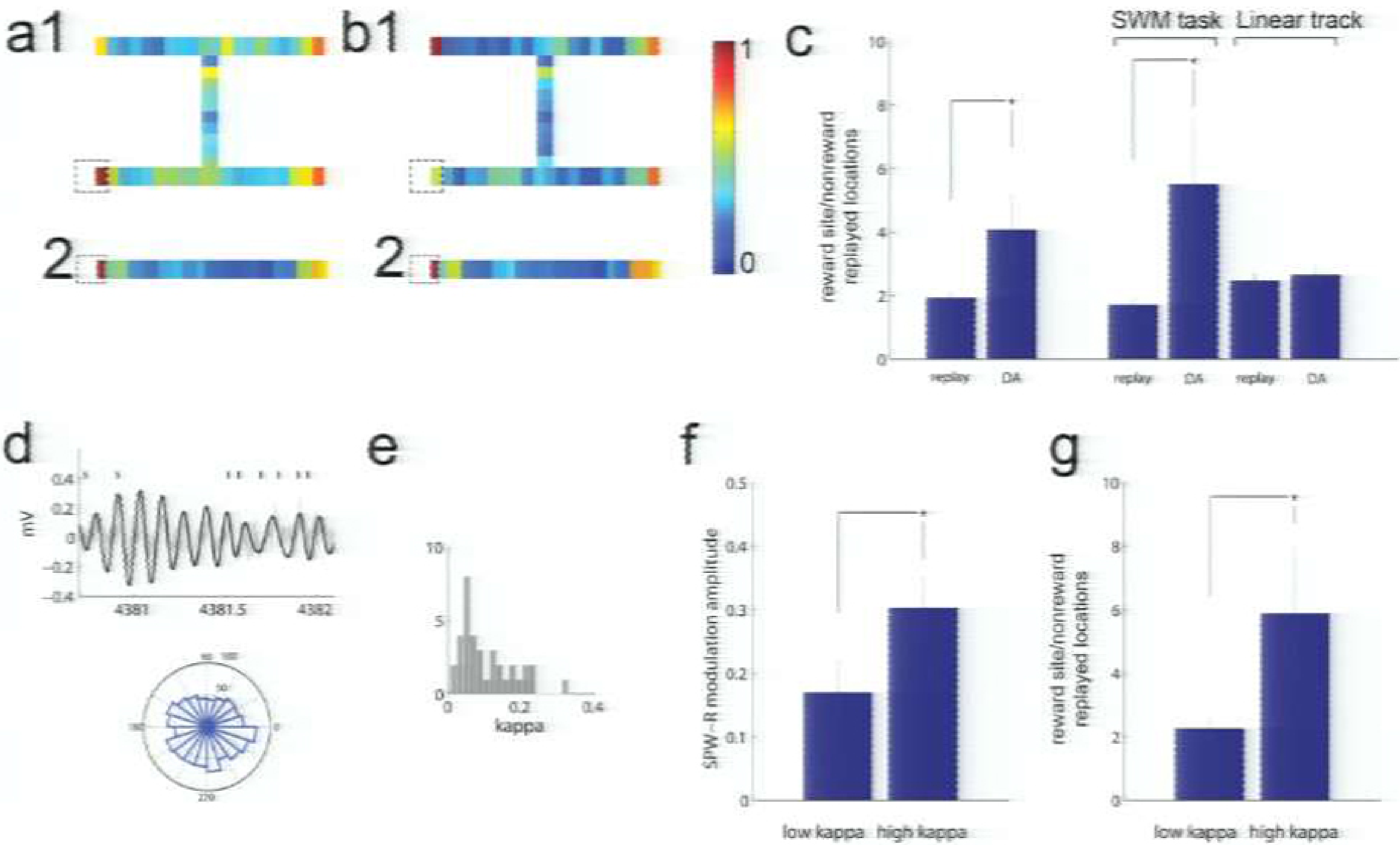


Dopamine unit modulation at hippocampal SPW-R bursts depends on replay content

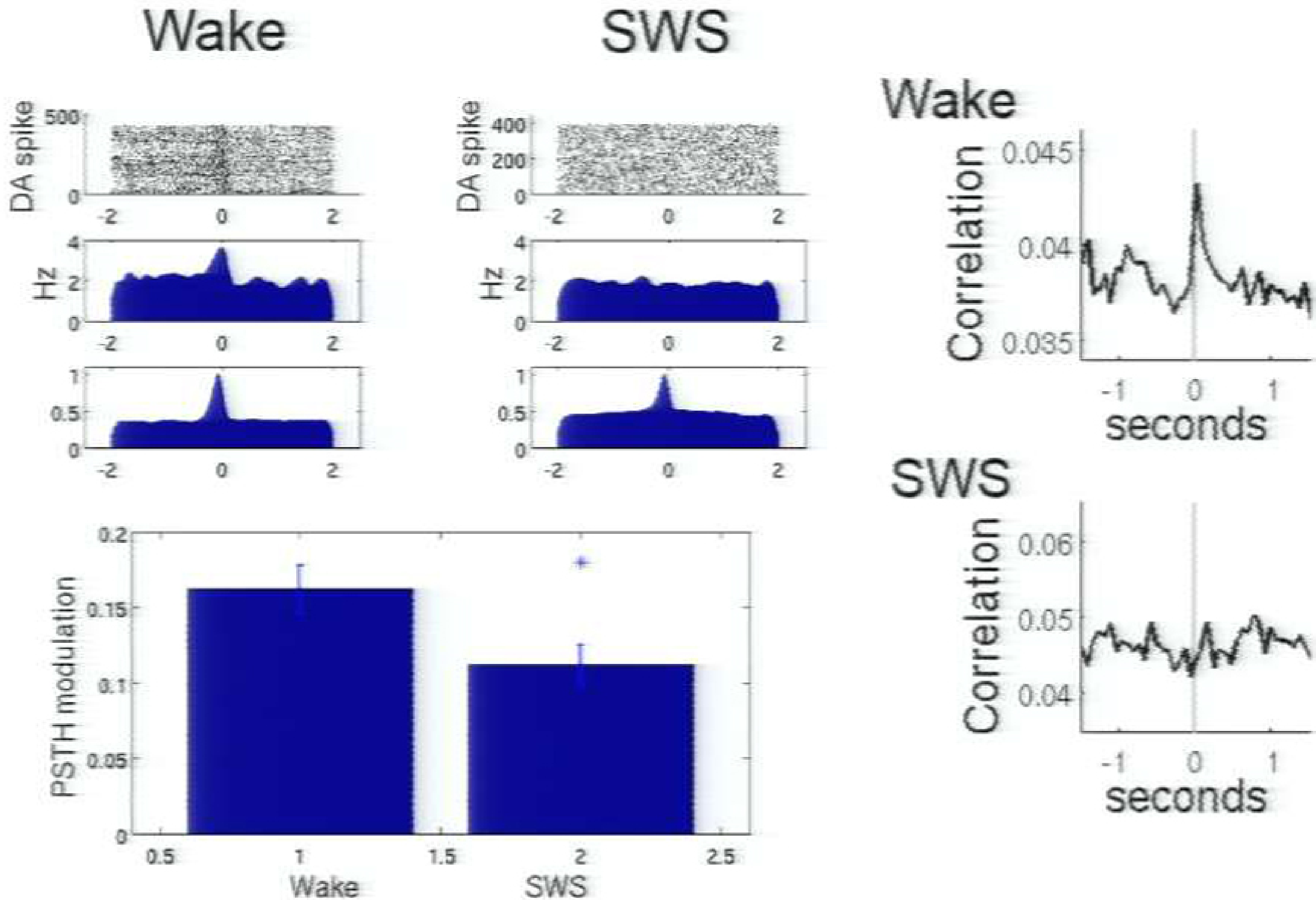
replay Non-replay



Dopamine units preferentially coordinate with replay of reward site locations on the spatial working memory task and phase lock to hippocampal theta



Modulation of DA activity at SPW-Rs is reduced in slow wave sleep



Summary

- DA unit activity increases during trajectories to rewards, differentially represents correct over error trials, and correlates with Q-TD prediction error in a spatial task.
- Hippocampal SPW-R bursts are associated with the modulation of DA units.
- Hippocampal theta phase-locking of DA unit activity correlates with the degree of SPW-R associated modulation.
- DA coordination with SPW-R bursts depends on replay content:
  - Replay of spatial sequences is associated with greater modulation.
  - DA units preferentially relate to replay of reward locations.
- SPW-R modulation of DA units is reduced in slow wave sleep.

Overall summary

- Sequence memory can be encoded in the hippocampus during active behavior.
- Sequence memory is subsequently replayed during sleep in both the hippocampus and neocortex.
- The content of reactivated memory during sleep can be biased by external manipulation.
- Sequence memory replayed during quiet wakefulness is associated reward information and may serve a different role in learning than replay during sleep.

Wilson Lab present and former

Albert Lee (Janelia Farm)	Non-REM replay
Daoyun Ji (Baylor)	H-Visual cortex
David Foster (J. Hopkins)	Awake replay
Fabian Kloosterman (Leuven)	Extended awake replay
Tom Davidson (Stanford)	Extended awake replay
Dan Bendor (UCL)	Biased sleep replay
Steve Gomperts (Harvard)	VTA and reward

Overall summary

- Sequence memory can be encoded in the hippocampus during active behavior.
- Sequence memory is subsequently replayed during sleep in both the hippocampus and neocortex.
- The content of reactivated memory during sleep can be biased by external manipulation.
- Sequence memory replayed during quiet wakefulness is associated reward information and may serve a different role in learning than replay during sleep.

Wilson Lab present and former

Albert Lee (Janelia Farm)	Non-REM replay
Daoyun Ji (Baylor)	H-Visual cortex
David Foster (J. Hopkins)	Awake replay
Fabian Kloosterman (Leuven)	Extended awake replay
Tom Davidson (Stanford)	Extended awake replay
Dan Bendor (UCL)	Biased sleep replay
Steve Gomperts (Harvard)	VTA and reward

NIPS Thanks Its Sponsors

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

amazon.com

Microsoft Research

Google

facebook

SKYTREE THE MACHINE LEARNING COMPANY

TWO SIGMA

United Technologies Research Center

D E Shaw & Co

YAHOO! LABS

IBM Research

xerox

TOYOTA

DRW TRADING GROUP

millionshort

criteo

PDT PARTNERS

Springer Machine Learning Journal

Disney Research

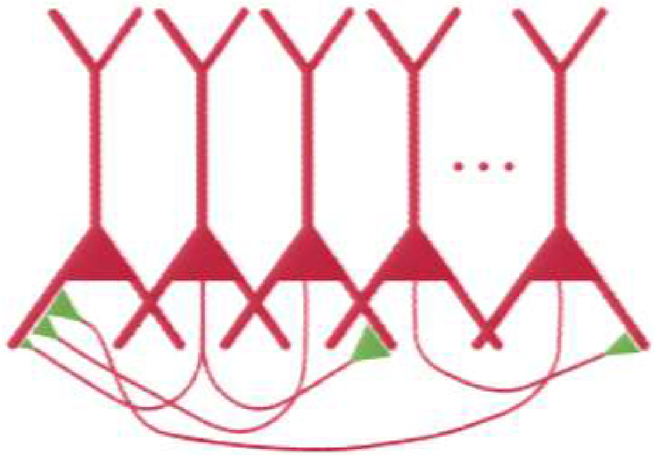
Correlations strike back (again): the case of associative memory retrieval

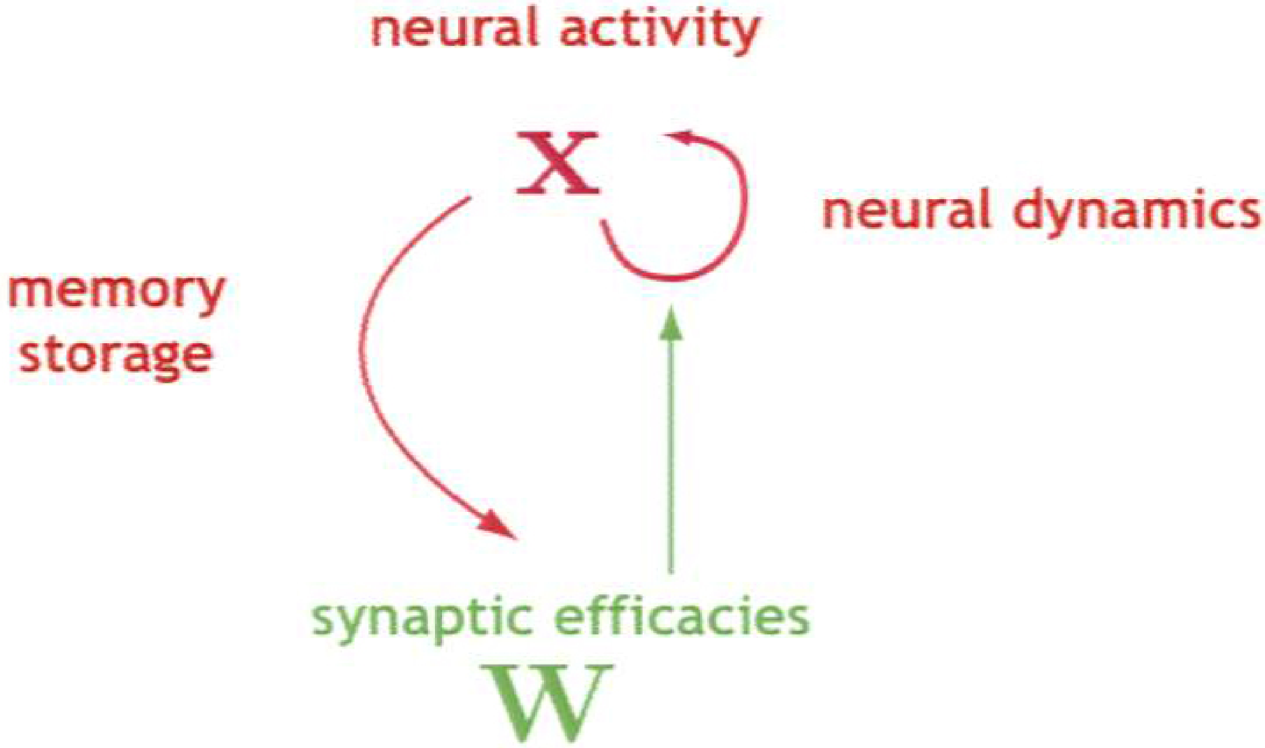
Cristina Savin

CBL, University of Cambridge, UK

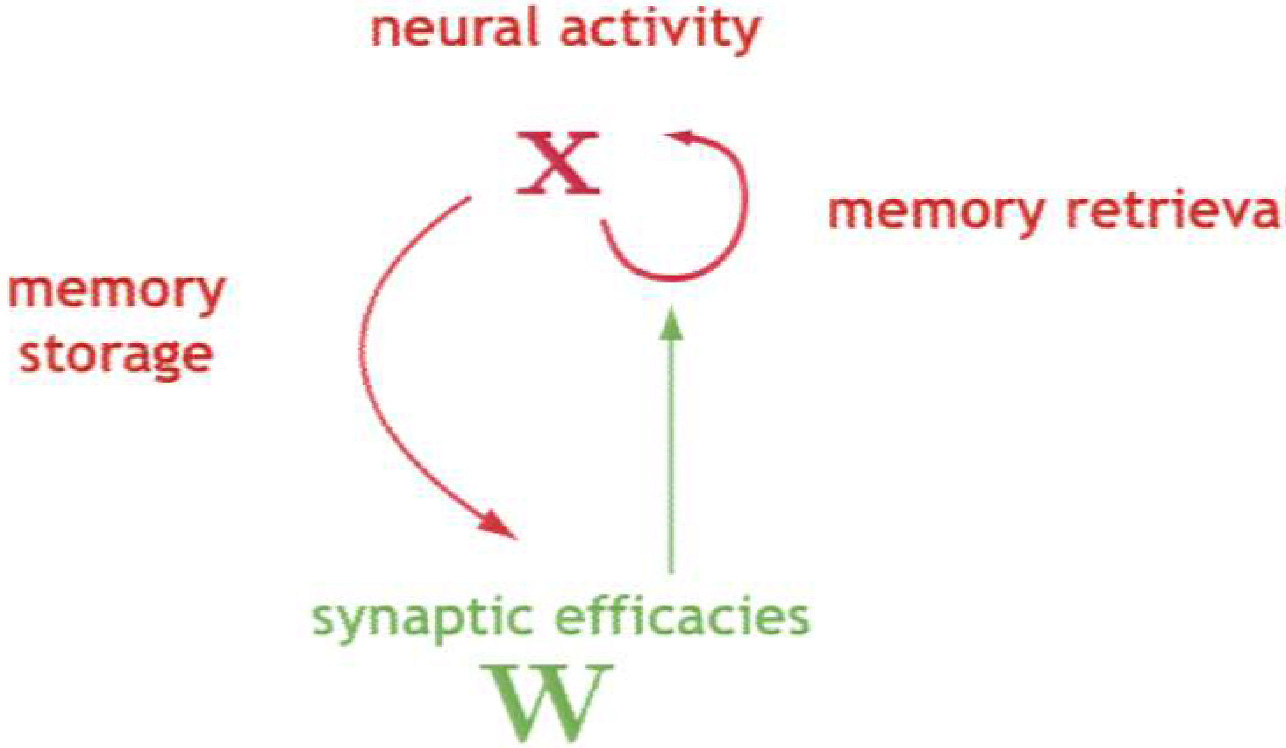
IST Austria

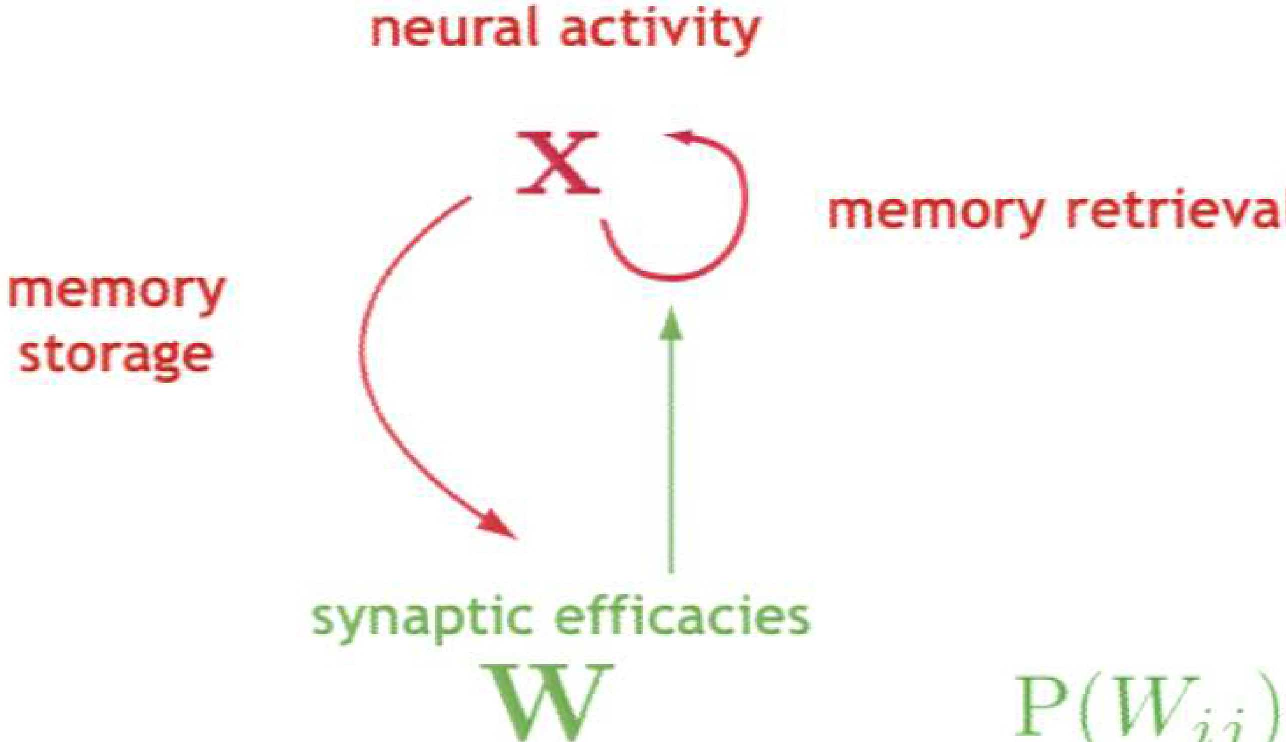
with Peter Dayan and Máté Lengyel





neural activity
x
neural dynamics
memory
storage
synaptic efficacies
W

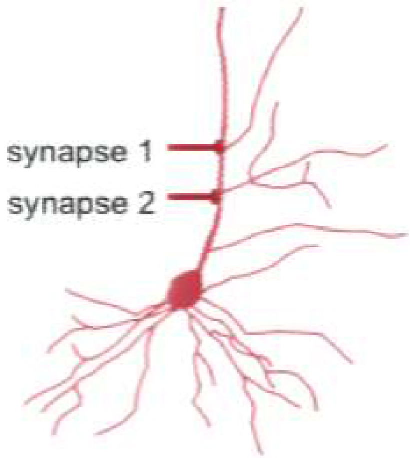

neural activity
x
memory retrieval
memory
storage
synaptic efficacies
W



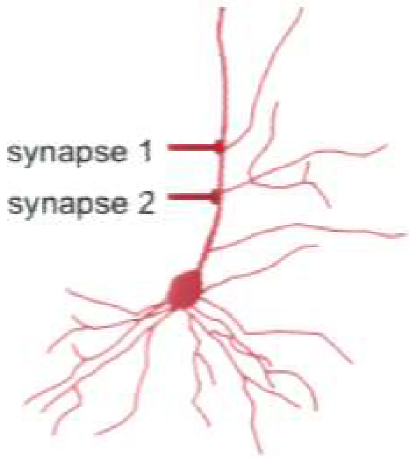
$$\mathrm{P}(W_{ij})$$

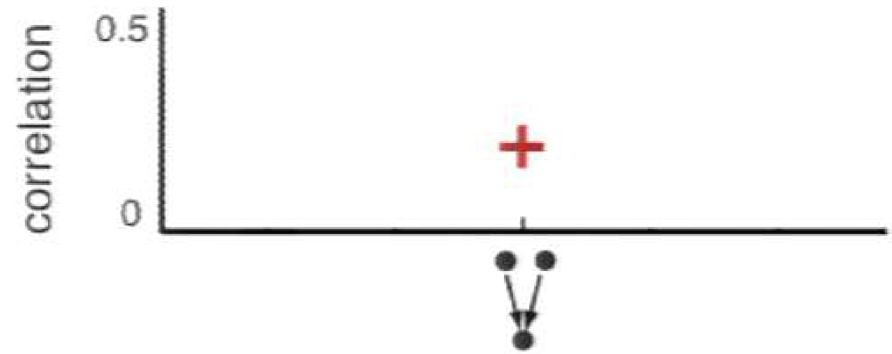
$$\mathrm{P}(W_{ij}, W_{kl})$$

Synaptic correlations in the cortex



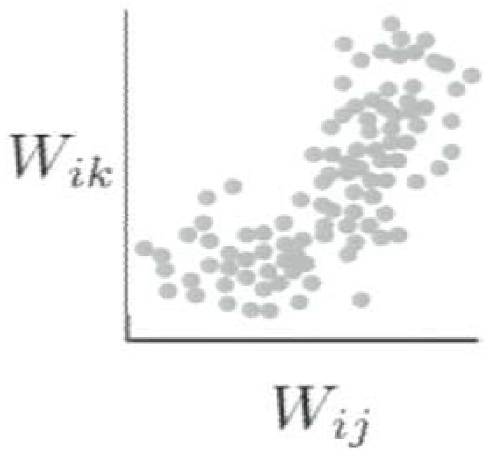
Synaptic correlations in the cortex



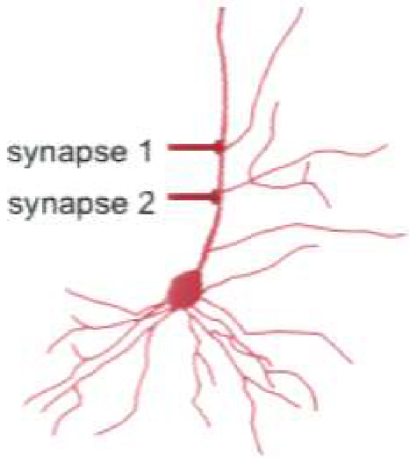


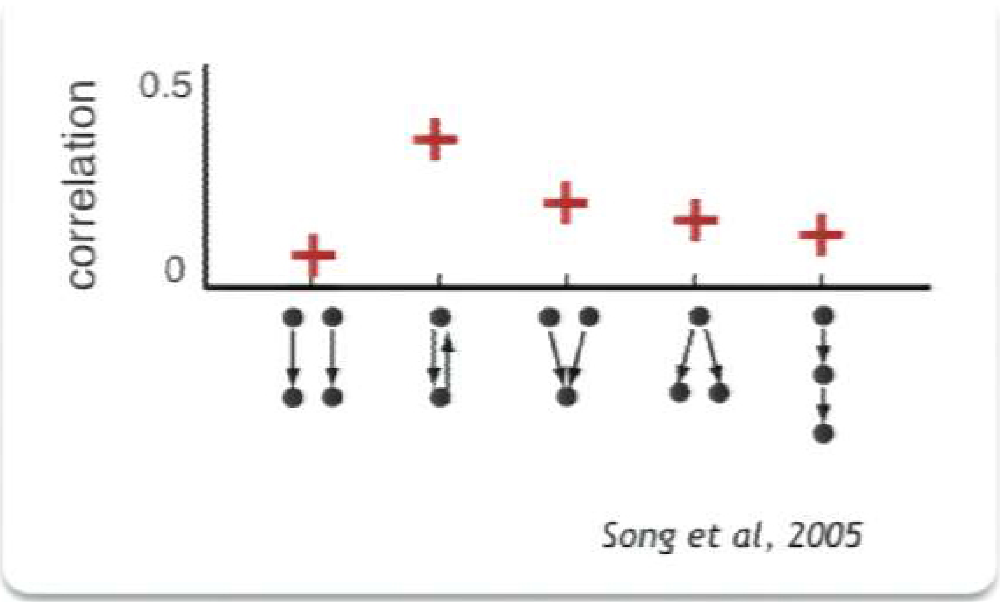
Song et al, 2005

$$P(W_{ij}, W_{ik})$$

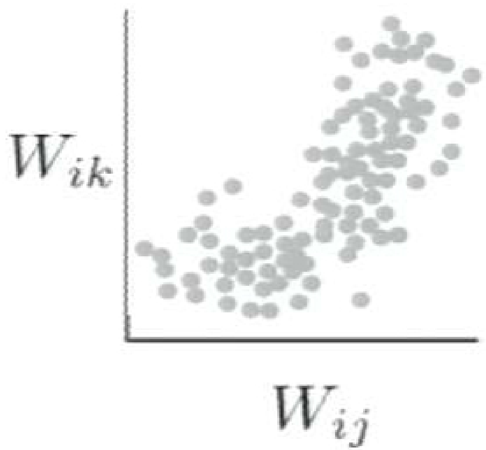


Synaptic correlations in the cortex





$$\mathrm{P}(W_{ij}, W_{ik})$$



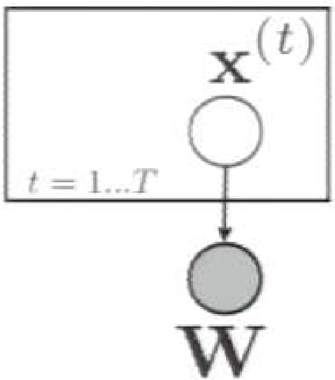
Where do they come from?

What do they mean for circuit function?

1. Where do synaptic correlations come from ?

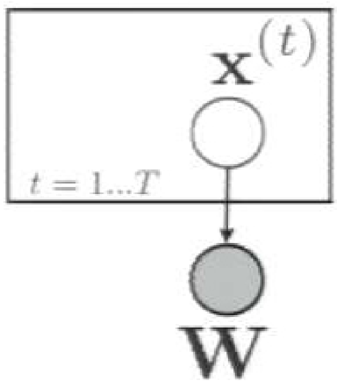
1. Where do synaptic correlations come from ?

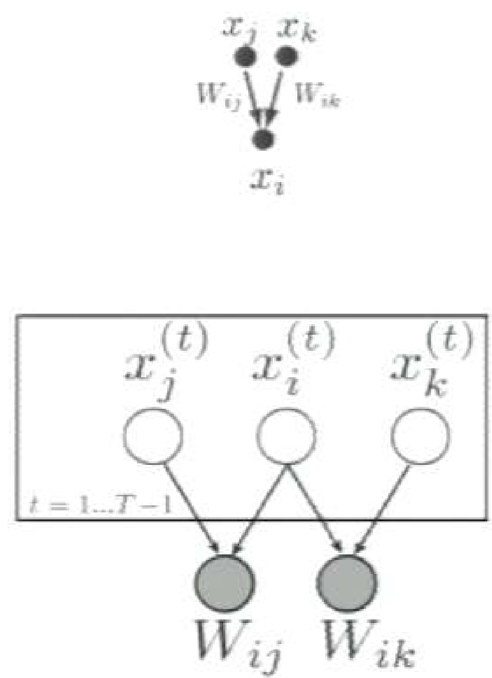
$$\mathrm{P}(\mathbf{W})$$



1. Where do synaptic correlations come from ?

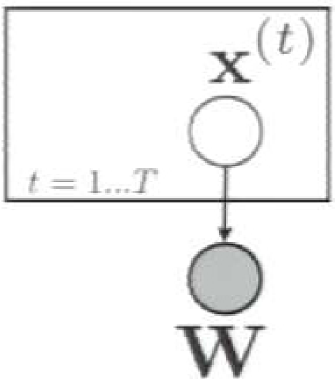
$P(\mathbf{W})$

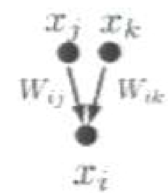




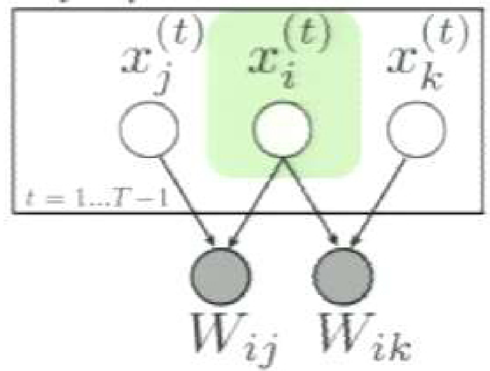
1. Where do synaptic correlations come from ?

$$P(\mathbf{W})$$



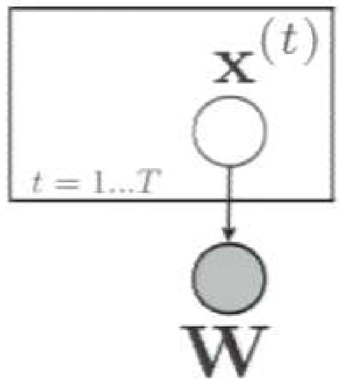


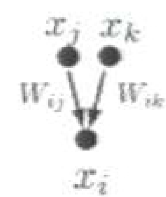
shared source of variability
for synapses on the same cell



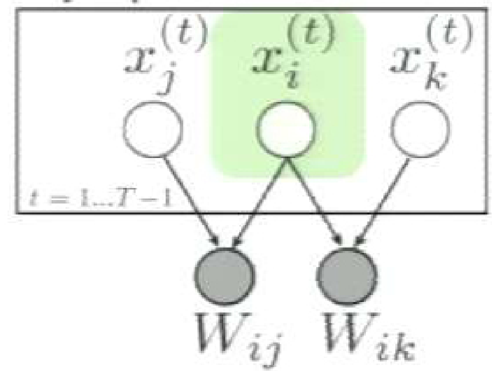
1. Where do synaptic correlations come from ?

$$P(\mathbf{W})$$





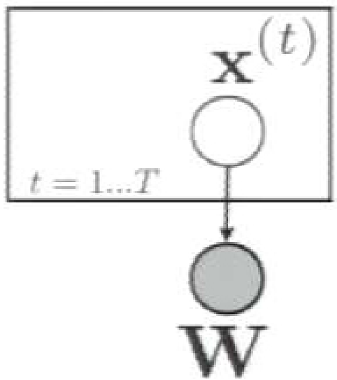
shared source of variability
for synapses on the same cell

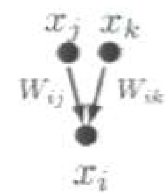


dependencies between synapses
sharing a pre- or post- synaptic partner

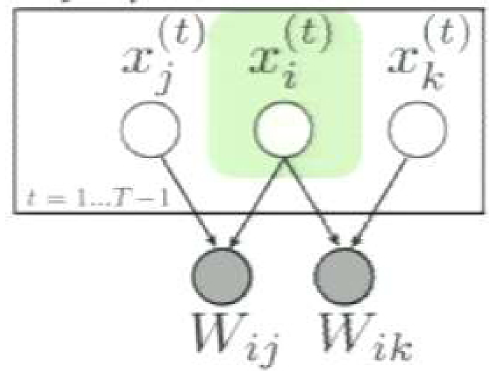
1. Where do synaptic correlations come from ?

$$P(\mathbf{W})$$





shared source of variability
for synapses on the same cell

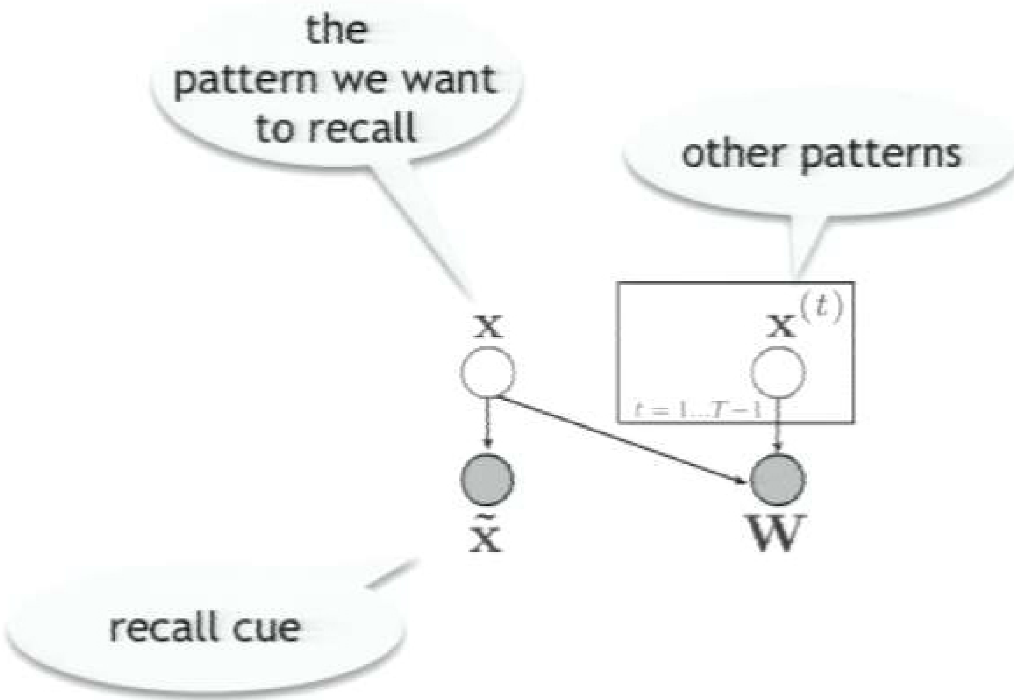


dependencies between synapses
sharing a pre- or post- synaptic partner

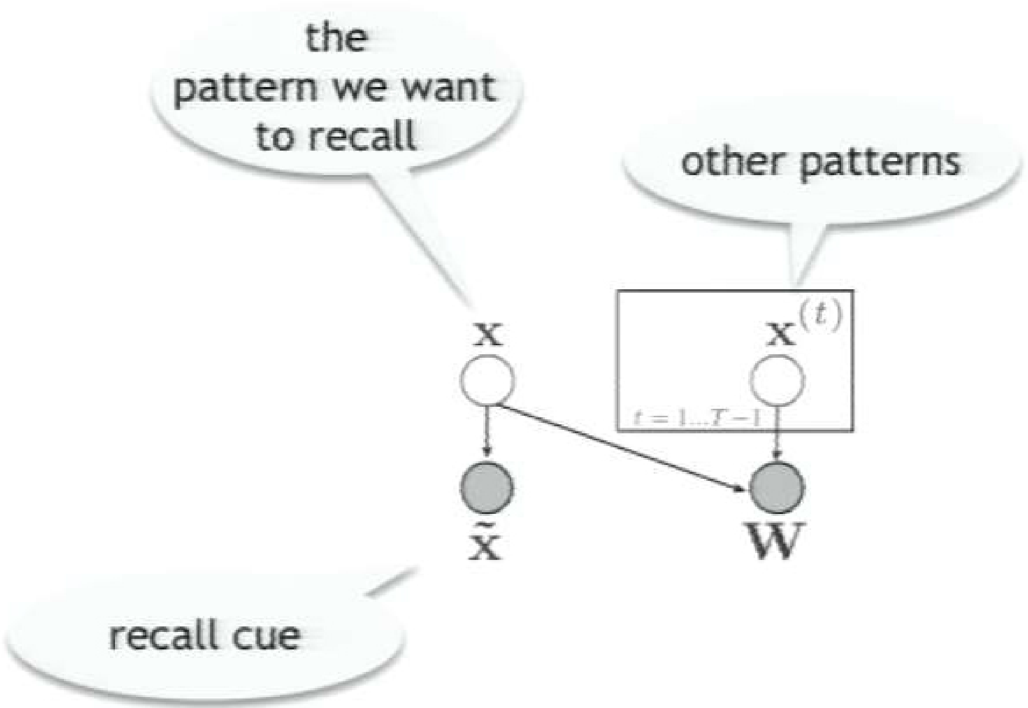
synaptic correlations
are a natural consequence
of synaptic plasticity

2. What do they mean for memory recall?

2. What do they mean for memory recall?



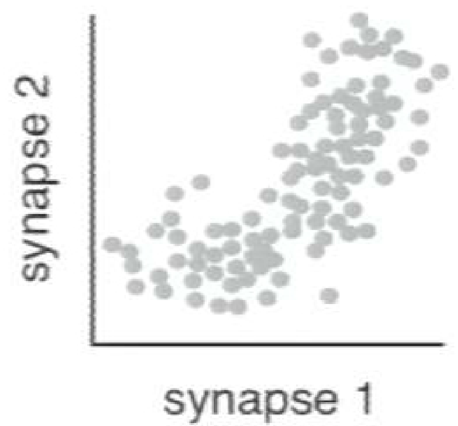
2. What do they mean for memory recall?



memory recall: $$\mathrm{P}(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto \mathrm{P}(\mathbf{x})\mathrm{P}(\tilde{\mathbf{x}}|\mathbf{x})\mathrm{P}(\mathbf{W}|\mathbf{x})$$

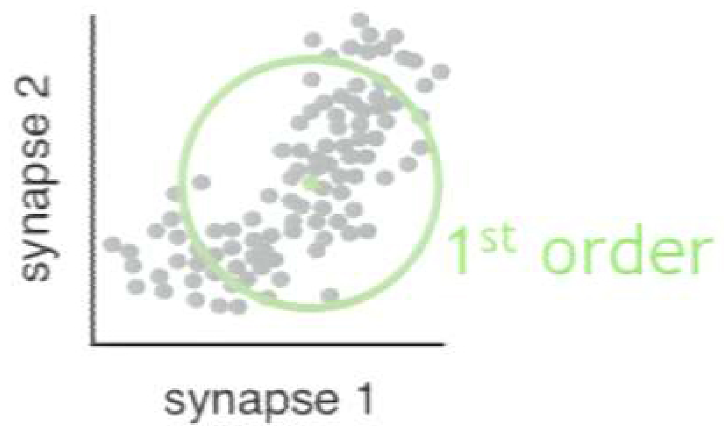
$$P(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$

$$P(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$



intractable

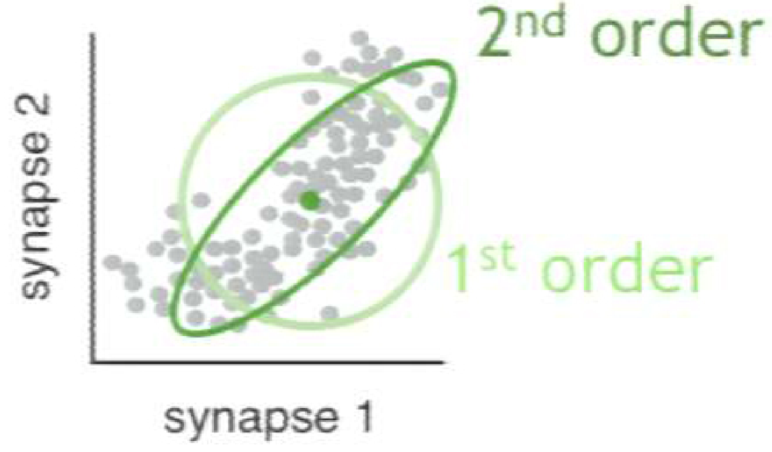
$$P(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$



approximation: maximum entropy distribution

intractable

$$P(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$

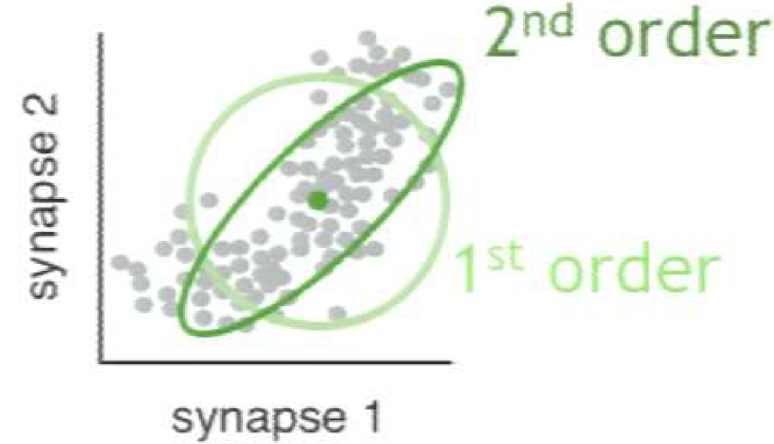


approximation: maximum entropy distribution

intractable

$$P(\mathbf{x}|\mathbf{W},\tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$

Can we improve recall performance if we take into account synaptic correlations?

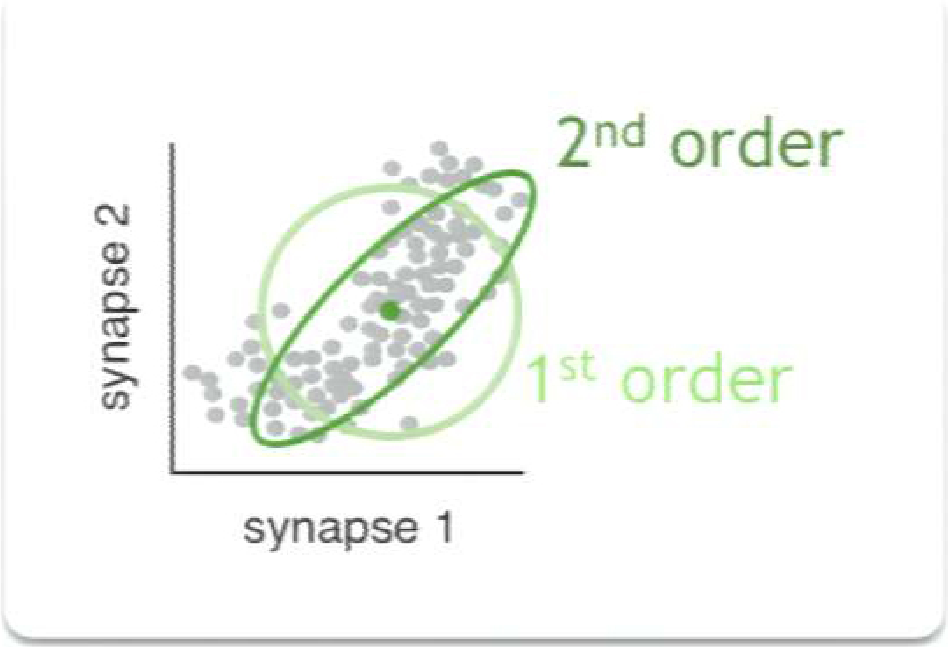


approximation: maximum entropy distribution

intractable

$$P(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$

Can we improve recall performance if we take into account synaptic correlations?

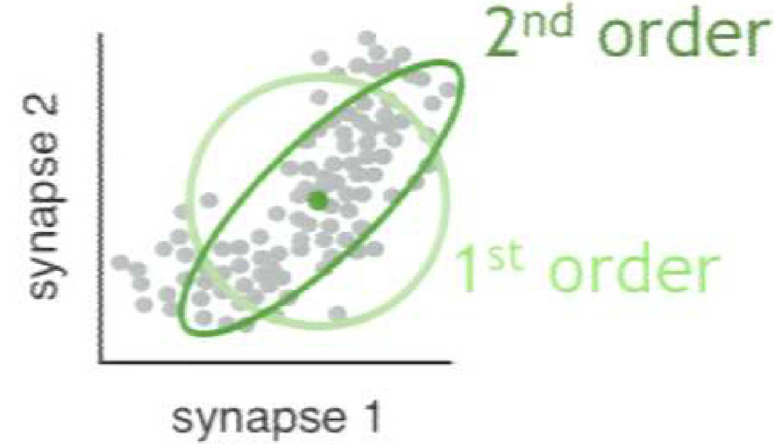


approximation: maximum entropy distribution

intractable

$$P(\mathbf{x}|\mathbf{W}, \tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$

Can we improve recall performance if we take into account synaptic correlations?



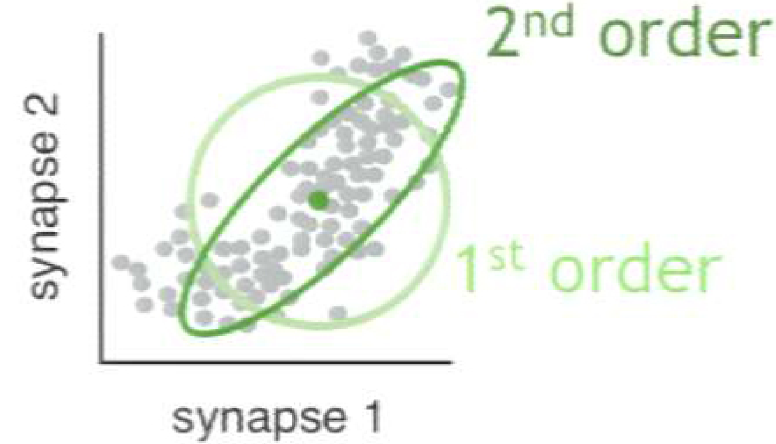
approximation: maximum entropy distribution

intractable

$$P(\mathbf{x}|\mathbf{W},\tilde{\mathbf{x}}) \propto P(\mathbf{x})P(\tilde{\mathbf{x}}|\mathbf{x})P(\mathbf{W}|\mathbf{x})$$

neural dynamics

Can we improve recall performance if we take into account synaptic correlations?

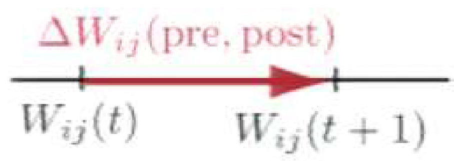


approximation: maximum entropy distribution

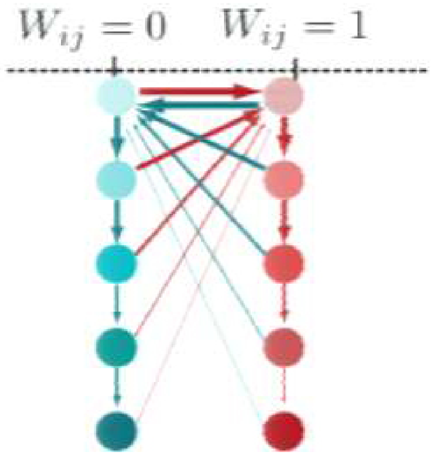
plasticity
rule

plasticity rule

additive



bounded synapses



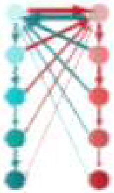
Fusi et al 2005

plasticity
rule

additive

$\Delta W_{ij}(\text{pre}, \text{post})$

bounded
synapses



plasticity rule

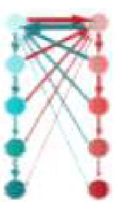
covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1

additive

$\Delta W_{ij}(\text{pre}, \text{post})$

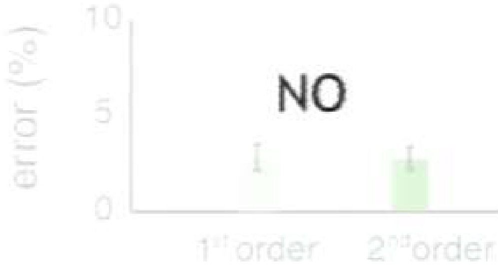
bounded synapses

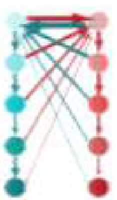


	plasticity rule	2nd order better than 1st?	neural implementation
additive $\Delta W_{ij}(\text{pre}, \text{post})$	covariance rule	NO	simple linear (Hopfield)
bounded synapses			

covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1





	plasticity rule	2^{nd} order better than 1^{st}?	neural implementation

additive

$\Delta W_{ij}(\text{pre}, \text{post})$

covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1

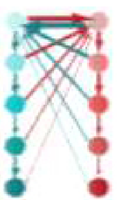
simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

NO

simple linear (Hopfield)

bounded synapses



plasticity rule	2nd order better than 1st?	neural implementation

additive

$\Delta W_{ij}(\text{pre}, \text{post})$

covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1

NO

simple linear (Hopfield)

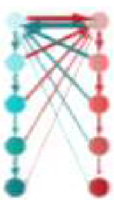
simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

YES

nonlinear inhibition

bounded synapses



	plasticity rule	2^{nd} order better than 1^{st}?	neural implementation
additive $\Delta W_{ij}(\text{pre}, \text{post})$	covariance rule	NO	simple linear (Hopfield)
	simple Hebb	YES	nonlinear inhibition
bounded synapses	postsyn. gated		

covariance rule

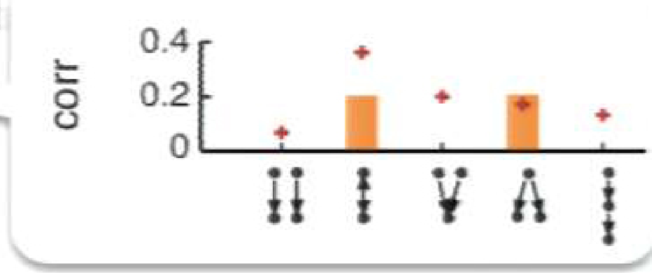
post \ pre	0	1
0	+1	-1
1	-1	+1

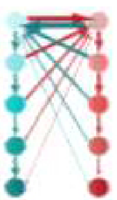
simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

postsyn. gated

post \ pre	0	1
0	-	-
1	D	P





	plasticity rule	2nd order better than 1st?	neural implementation
additive $\Delta W_{ij}(\text{pre}, \text{post})$	covariance rule	NO	simple linear (Hopfield)
	simple Hebb	YES	nonlinear inhibition
bounded synapses	postsyn. gated		

covariance rule

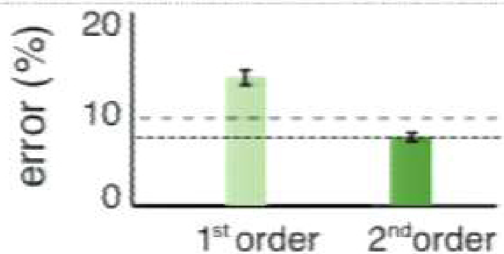
post \ pre	0	1
0	+1	-1
1	-1	+1

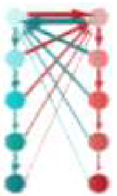
simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

postsyn. gated

post \ pre	0	1
0	-	-
1	D	P





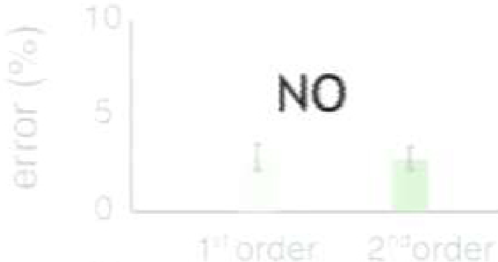
	plasticity rule	2nd order better than 1st?	neural implementation

additive

$\Delta W_{ij}(\mathrm{pre}, \mathrm{post})$

covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1



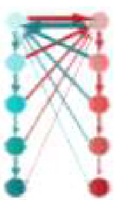
simple linear (Hopfield)

simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

nonlinear inhibition

bounded synapses



postsyn. gated

post \ pre	0	1
0	-	-
1	D	P

nonlinear dendritic integration

	plasticity rule	2^{nd} order better than 1^{st}?	neural implementation
additive $\Delta W_{ij}(\text{pre, post})$	covariance rule	NO	simple linear (Hopfield)
	simple Hebb	YES	nonlinear inhibition
bounded synapses	postsyn. gated		nonlinear dendritic integration

covariance rule

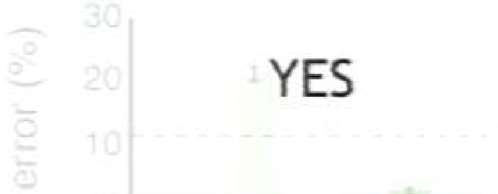
post \ pre	0	1
0	+1	-1
1	-1	+1

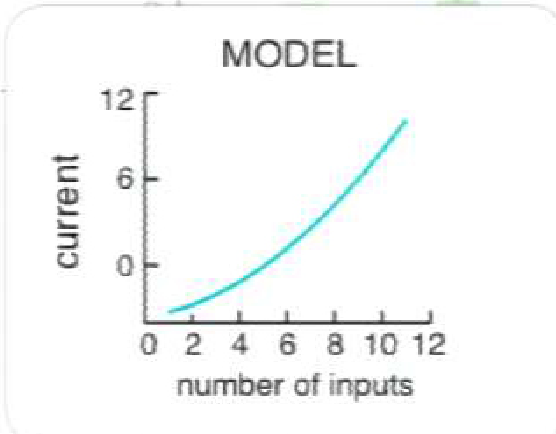
simple Hebb

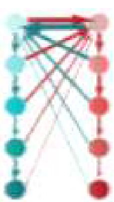
post \ pre	0	1
0	-	-
1	-	+1

postsyn. gated

post \ pre	0	1
0	-	-
1	D	P







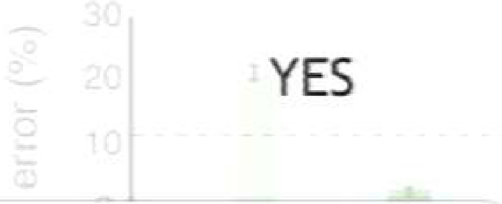
	plasticity rule	2^{nd} order better than 1^{st}?	neural implementation
additive	covariance rule	NO	simple linear (Hopfield)
$\Delta W_{ij}(\text{pre}, \text{post})$	simple Hebb	YES	nonlinear inhibition
			nonlinear dendritic integration

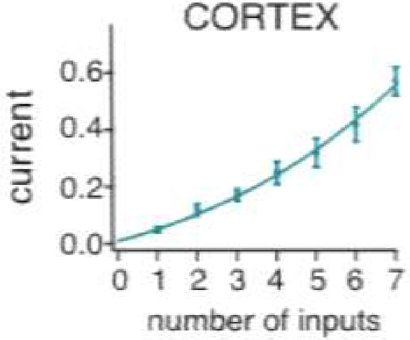
covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1

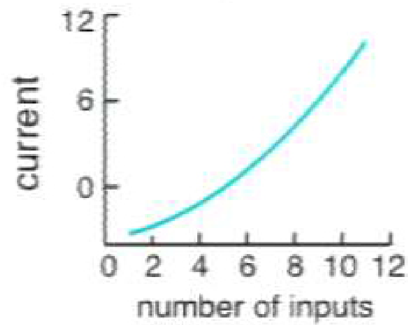
simple Hebb

post \ pre	0	1
0	-	-
1	-	+1



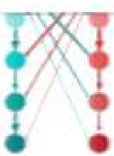


(Branco et al, 2011)



b

s



	plasticity rule	2nd order better than 1st?	neural implementation

additive

$\Delta W_{ij}(\text{pre}, \text{post})$

covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1

NO

simple linear (Hopfield)

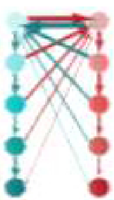
simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

YES

nonlinear inhibition

bounded synapses



postsyn. gated

post \ pre	0	1
0	-	-
1	D	P

YES

nonlinear dendritic integration

plasticity rule	2nd order better than 1st?	neural implementation

additive

$\Delta W_{ij}(\mathrm{pre}, \mathrm{post})$

covariance rule

post \ pre	0	1
0	+1	-1
1	-1	+1

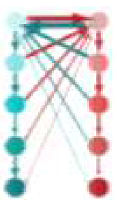
simple linear (Hopfield)

simple Hebb

post \ pre	0	1
0	-	-
1	-	+1

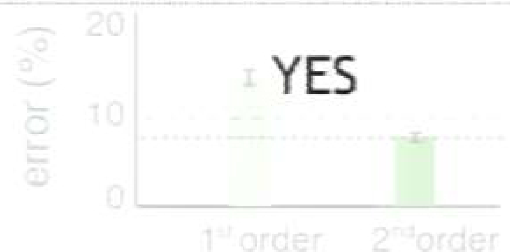
nonlinear inhibition

bounded synapses



postsyn. gated

post \ pre	0	1
0	-	-
1	D	P

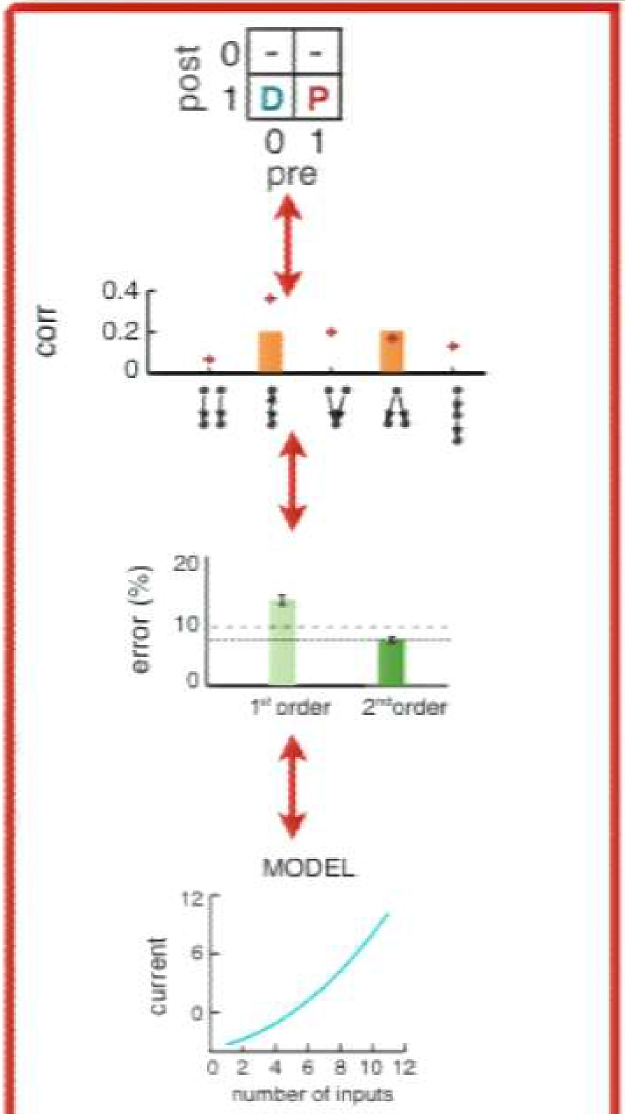
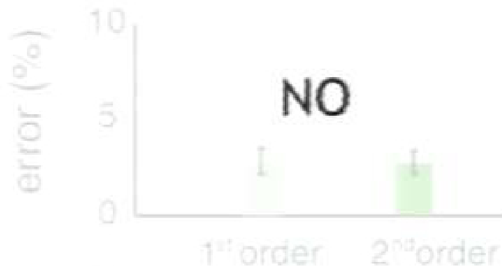


nonlinear dendritic integration

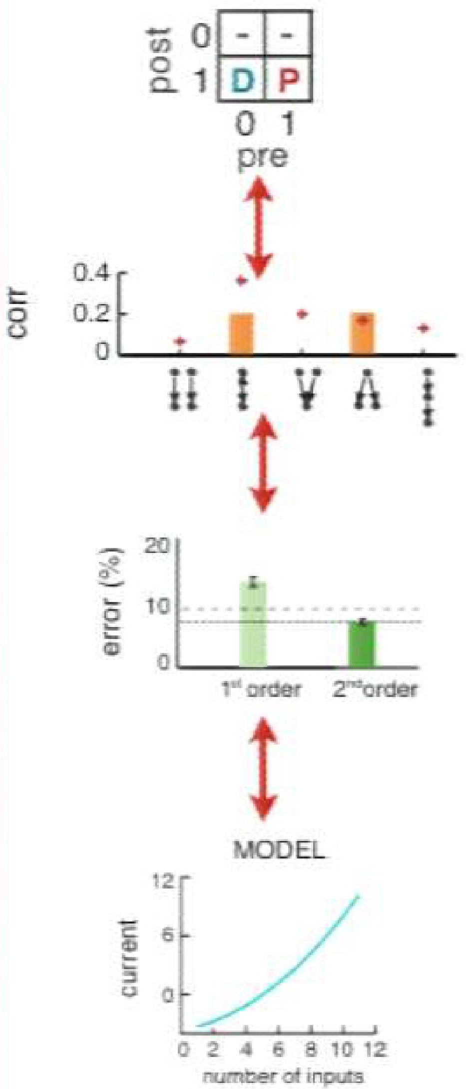
presyn. gated

post \ pre	0	1
0	-	D
1	-	P

linear dendritic integration

plasticity rule	2^{nd} order better than 1^{st}?	neural implementation
		simple linear (Hopfield)
		nonlinear inhibition
		nonlinear dendritic integration
		linear dendritic integration

plasticity rule	2nd order better than 1st?	neural implementation



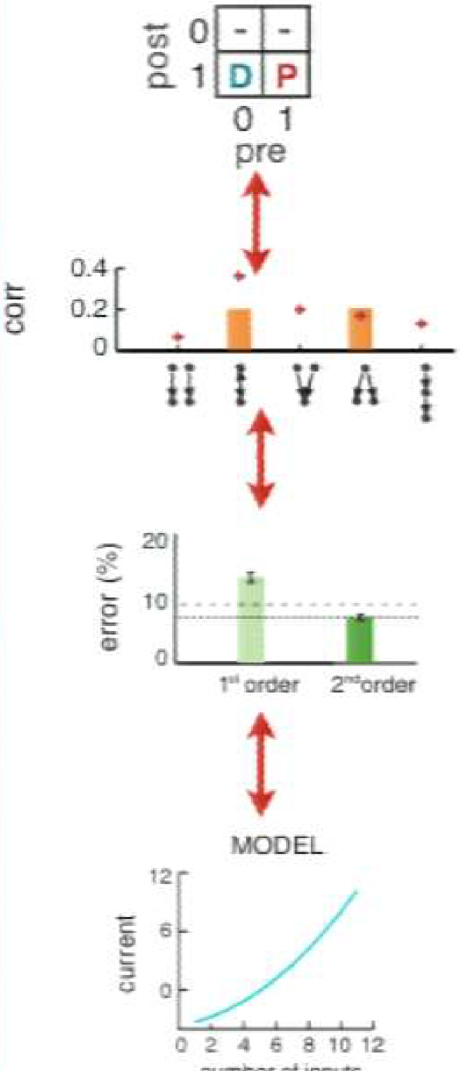
synaptic correlations need to be taken into account for recall

simple linear (Hopfield)

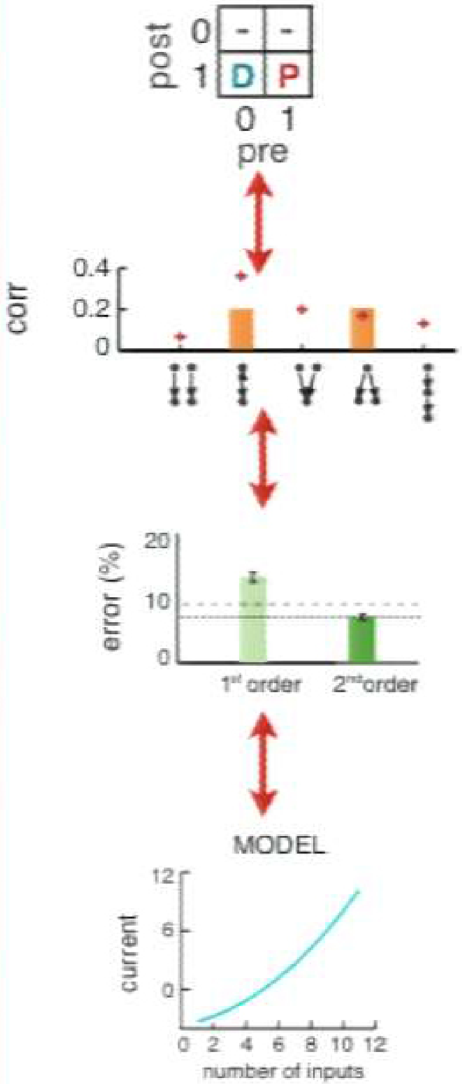
nonlinear inhibition

nonlinear dendritic integration

linear dendritic integration

plasticity rule	2^{nd} order better than 1^{st}?	neural implementation
	synaptic correlations need to be taken into account for recall	circuit nonlinearities can be understood as adaptations for approx. optimal recall

plasticity rule | 2^{nd} order better than 1^{st}? | neural implementation



synaptic correlations need to be taken into account for recall

circuit nonlinearities can be understood as adaptations for approx. optimal recall

For technical details: poster Fri61

NIPS Thanks Its Sponsors

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

amazon.com

Microsoft Research

Google

facebook

SKYTREE
THE MACHINE LEARNING COMPANY

TWO SIGMA

United Technologies Research Center

D E Shaw & Co

YAHOO! LABS

IBM Research

xerox

DRW TRADING GROUP

TOYOTA

millionshort

criteo

PDT PARTNERS

Springer
Machine Learning Journal

Disney Research

A memory frontier for complex synapses

Surya Ganguli

Dept. of Applied Physics
and, by courtesy,
Neurobiology and Electrical Engineering

Stanford University

Joint work with: Subhaneil Lahiri

a.k.a.

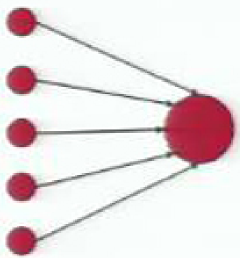
"Subhy"

Poster F50 Tonight

The synaptic basis for long-term memory storage

New York Times…

A gulf between theory and experiment



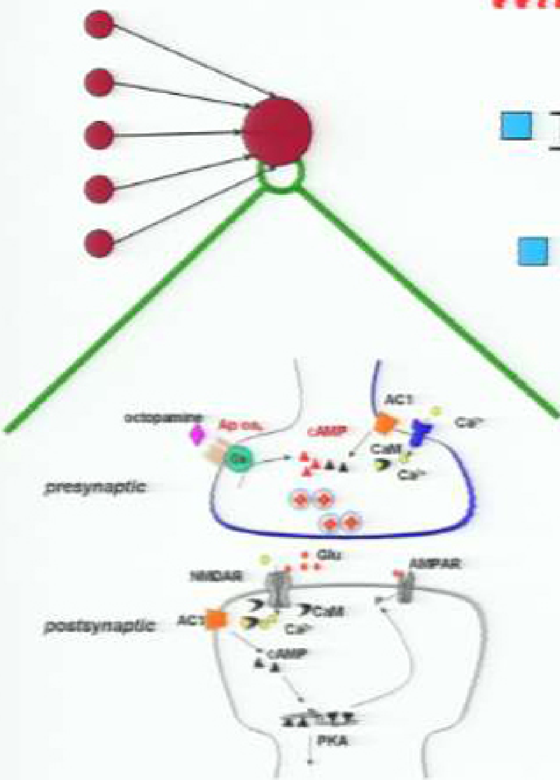
What is a synapse from neuron j to neuron i?

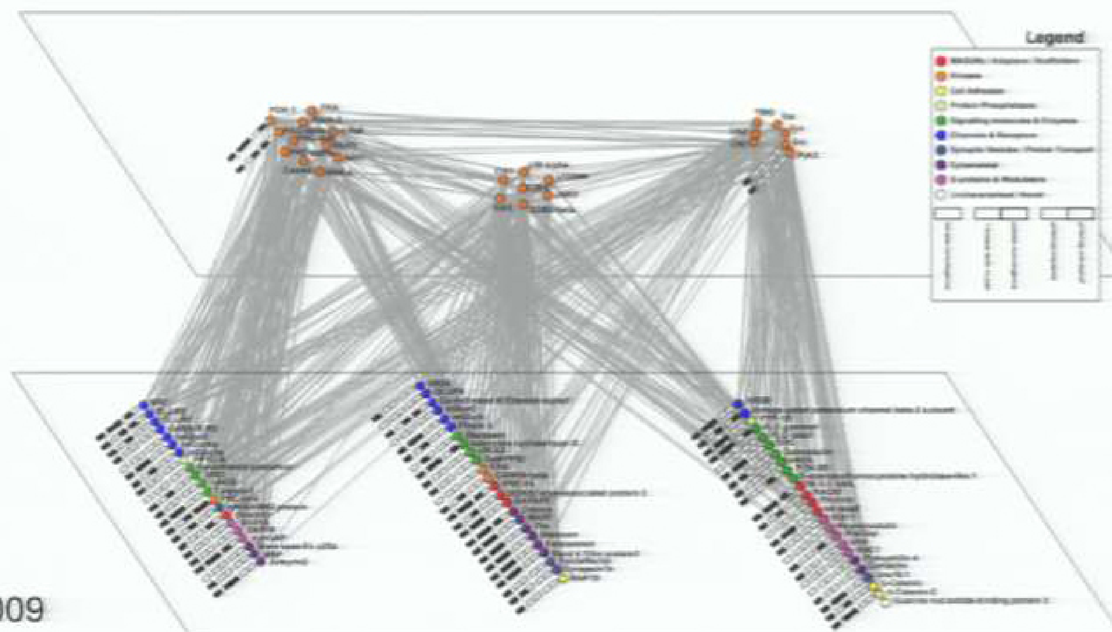
- <u>Theorist</u>: W_{ij} or J_{ij} ~ size of postsynaptic potential

A gulf between theory and experiment

What is a synapse from neuron j to neuron i?

- <u>Theorist</u>: W_{ij} or J_{ij} ~ size of postsynaptic potential
- <u>Experimentalist</u>: AMPA, NMDA, CAMKII, MAPK, CREB, MHC-I, second messengers, membrane protein regulatio intracellular trafficking, new protein synthesis ……….



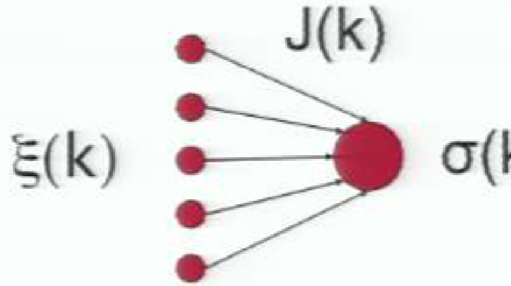


Coba et. al.
Science Signalling 2009

Memory capacity with scalar analog synapses

Consider the number of associations a neuron with N afferent synapses can store.

$$\sigma(k) = \mathrm{sgn}\,(J \cdot \xi(k) - \theta)$$



An online learning rule to store the desired association:

$$J(k+1) = e^{-1/\tau} J(k) + \sigma(k)\, \xi(k)$$

i.e. 1) Allows analog weights to decay slightly (forget the past inputs)
2) Add in the new association to the weight (learn a new input).

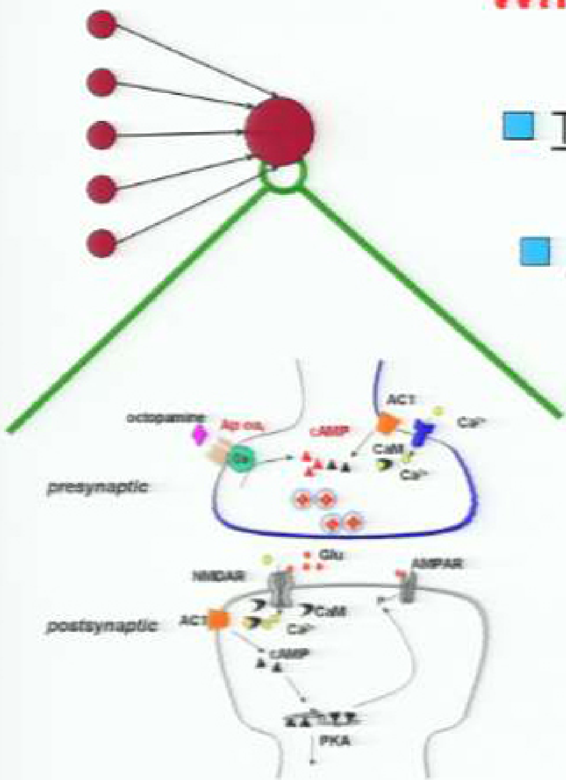
Memory capacity: How far back into the past can synapses reliably recall previously stored associations?

Answer: If τ is O(N) then the past O(N) associations can be recalled.

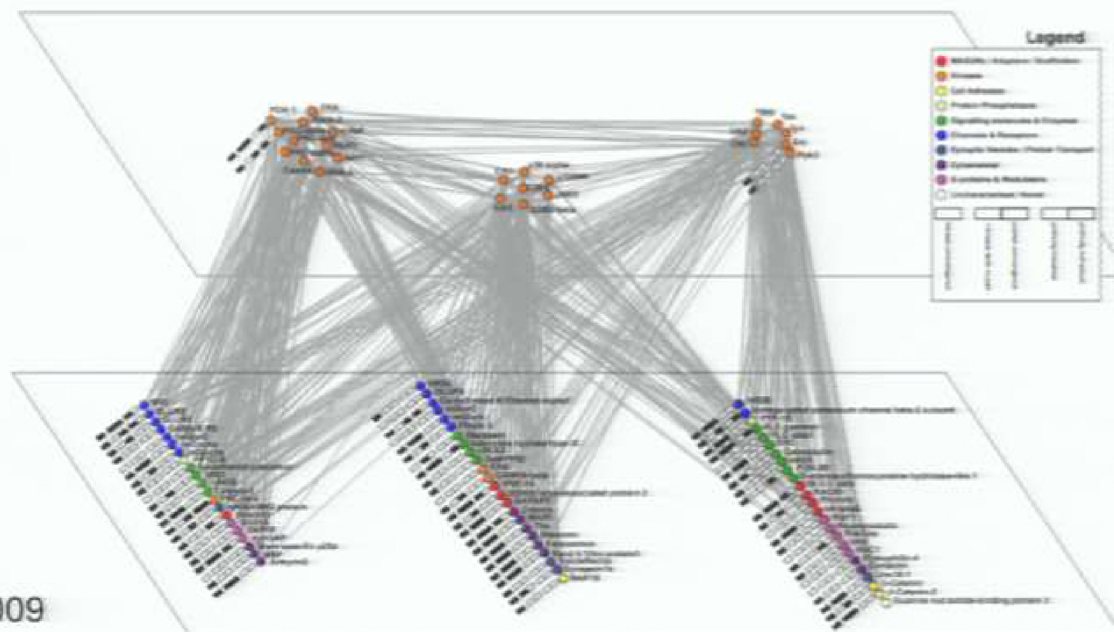
Problem: This solution relies on individual synapses to reliably maintain O(N) distinguishable analog states.

A gulf between theory and experiment

What is a synapse from neuron j to neuron i?



- Theorist: W_{ij} or J_{ij} ~ size of postsynaptic potential
- Experimentalist: AMPA, NMDA, CAMKII, MAPK, CREB, MHC-I, second messengers, membrane protein regulatio intracellular trafficking, new protein synthesis ……….

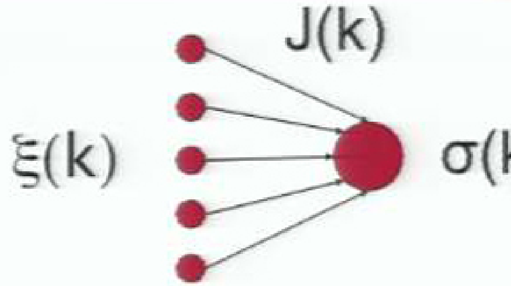


Coba et. al.
Science Signalling 2009

Memory capacity with scalar analog synapses

Consider the number of associations a neuron with N afferent synapses can store.

$$\sigma(k) = \mathrm{sgn}\,(J \cdot \xi(k) - \theta)$$



An online learning rule to store the desired association:

$$J(k+1) = e^{-1/\tau} J(k) + \sigma(k)\,\xi(k)$$

i.e. 1) Allows analog weights to decay slightly (forget the past inputs)
2) Add in the new association to the weight (learn a new input).

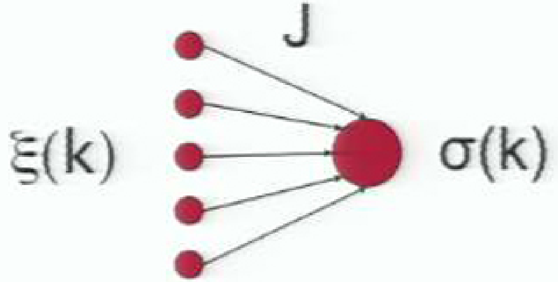
Memory capacity: How far back into the past can synapses reliably recall previously stored associations?

Answer: If τ is O(N) then the past O(N) associations can be recalled.

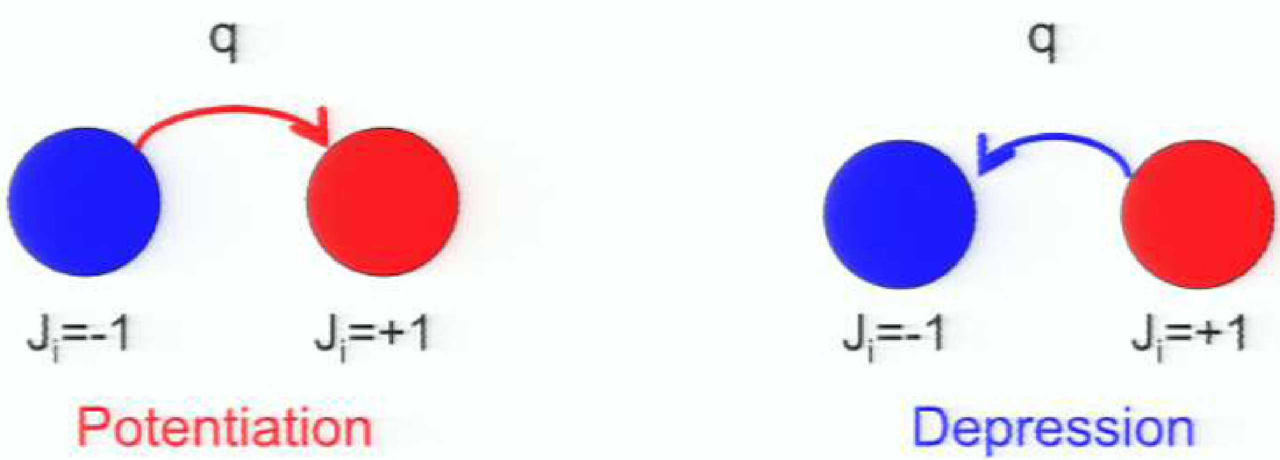
Problem: This solution relies on individual synapses to reliably maintain O(N) distinguishable analog states.

Memory capacity with binary synapses

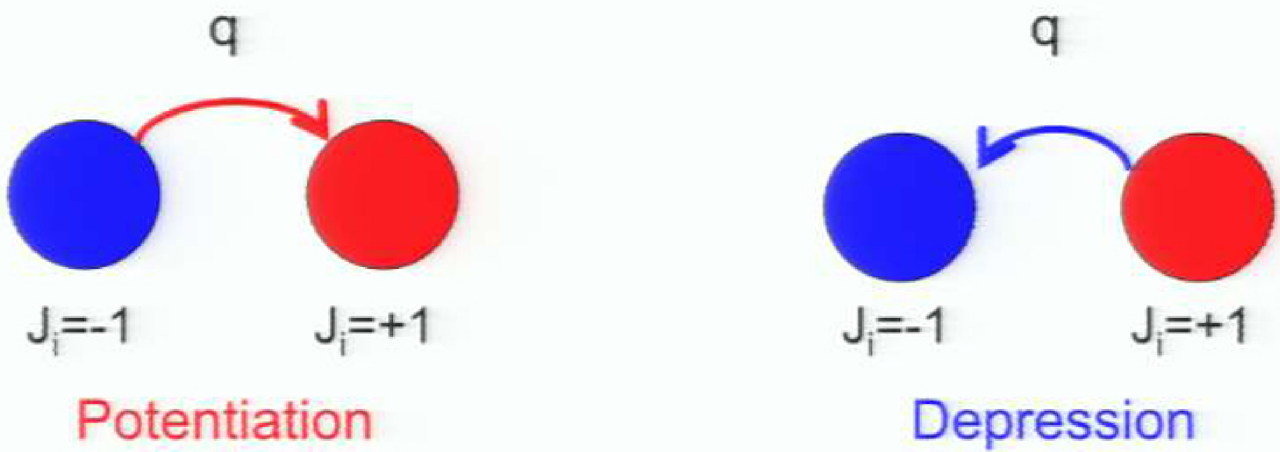
What about real synapses which can take only a finite number of distinguishable values for their strength?



For binary synapses each synapse $J_i = +1$ or -1. So you can no longer add an association to synaptic weights without running into boundaries.



Memory capacity with binary synapses



q = prob a synapse changes strength under appropriate conditions
N = number of synapses

Memory Capacity

$q = O(1)$	$\log N$	Quickly learn, quickly forget
$q = O(N^{-1/2})$	$N^{1/2}$	Sluggish to learn, slow to forget

Fusi and Amit 92

Synaptic complexity: from scalars to dynamical systems

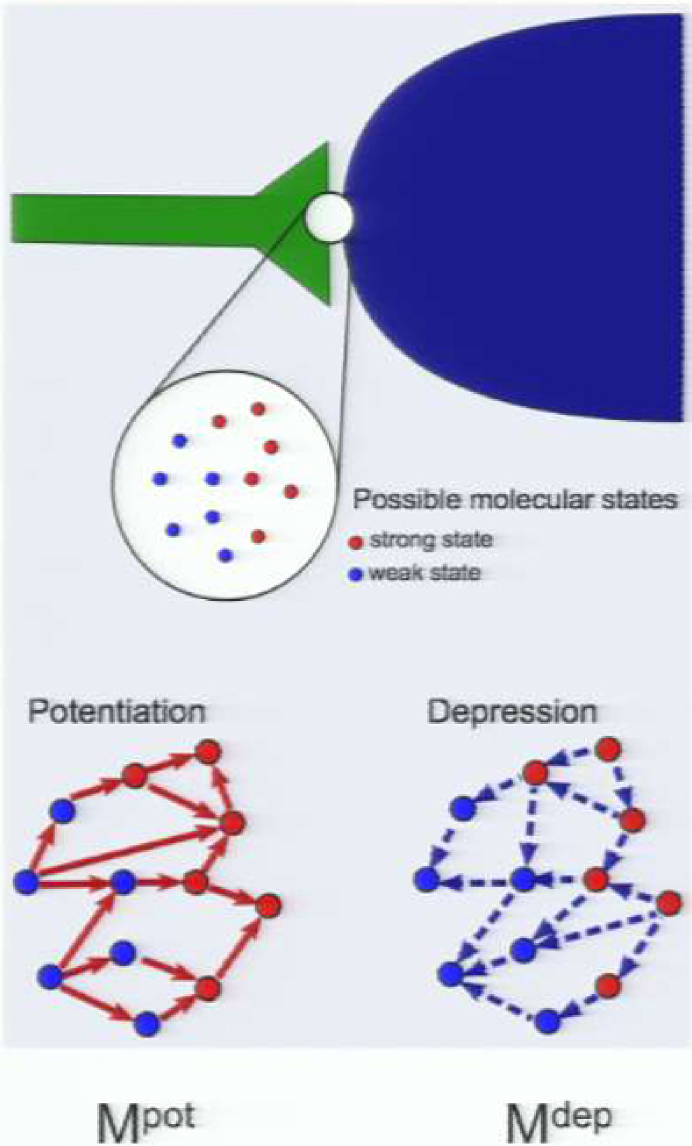
Experiment

Theory

We must expand our theoretical conception of a synapse from that of an simple scalar value to an entire (stochastic) dynamical system in its own right.

This yields a large universe of synaptic models to explore and understand.

Framework for synaptic dynamical systems

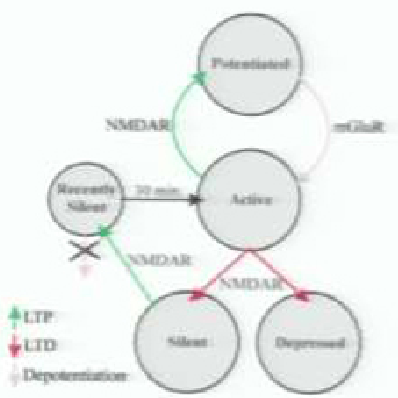


Theoretical approach:

A synapse is an arbitrary stochastic dynamical system with M internal states.

Some internal states correspond to a strong synapse, others a weak synapse.

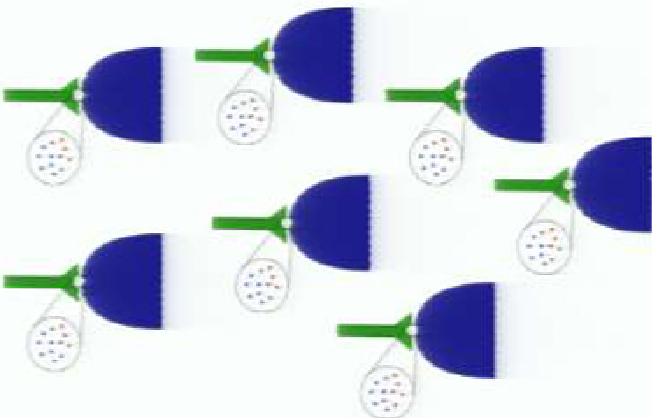
A candidate potentiation (depression) event induces an arbitrary stochastic transition between states.



Montgomery and Madison Neuron 2002

Ideal observer measure of memory capacity: SNR

A continuous stream of memories are stored (at poisson rate r) in a population of N synapses with M internal states.



The memory stored at time t=0 demands that some synapses potentiate, while others depress, yielding an ideal synaptic weight vector w_{ideal}.

The storage of future memories after t=0 changes the weight vector to w(t).

An upper bound on the quality of memory retrieval of any memory readout using neural activity is given by the SNR curve:

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

Each choice of

N, M, M^{pot} and M^{dep}

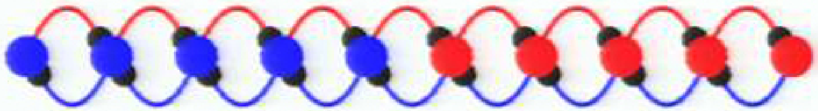
yields a different memory curve.

Fusi et. al. 2005, Fusi et. al. 2007, Barrett and van Rossum, 2008

Two example synaptic molecular networks

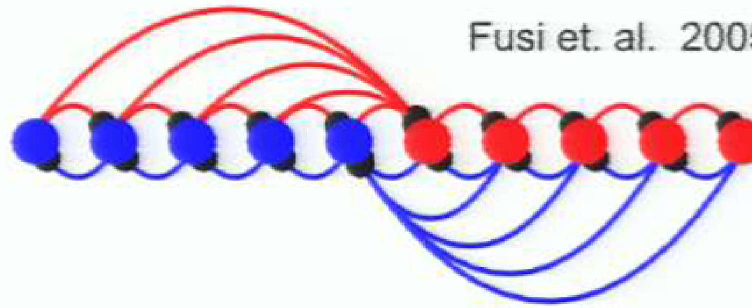
Serial Model

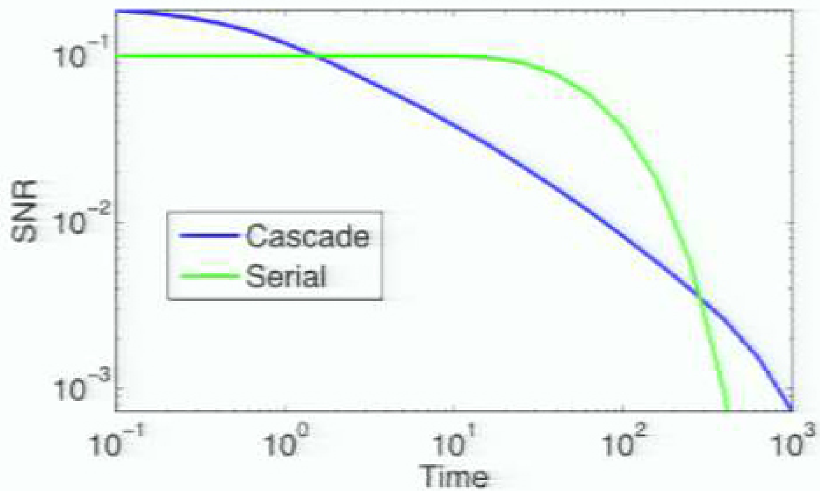
Leibold and Kempter 2008



Cascade Model

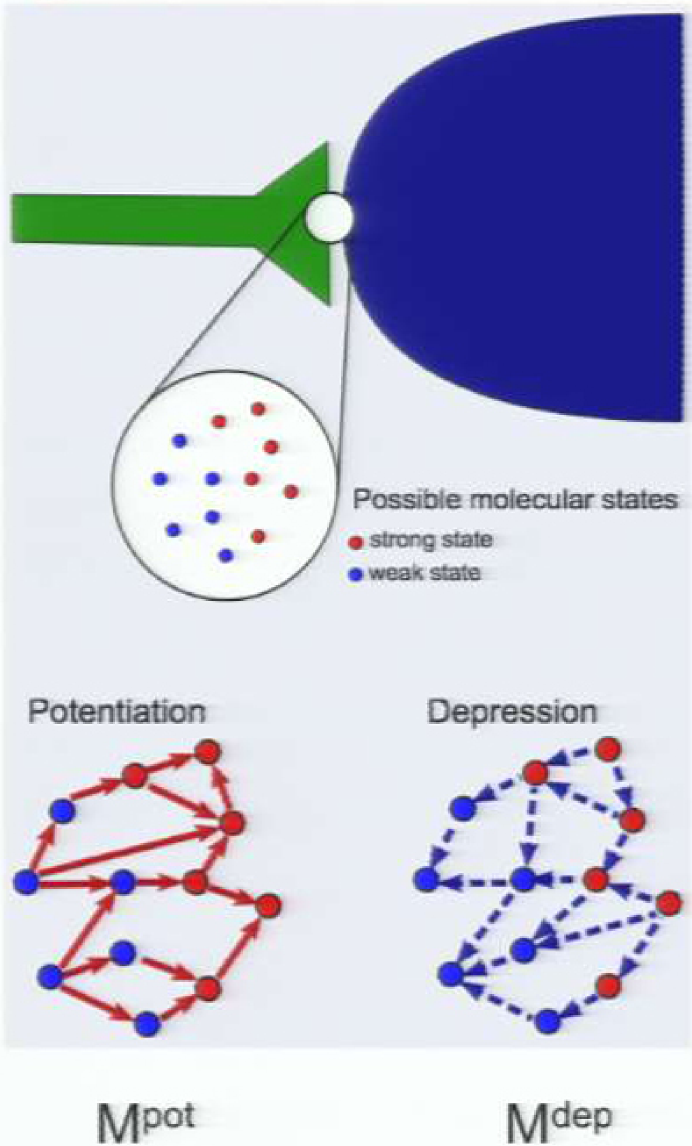
Fusi et. al. 200





To elucidate the functional contribution of molecular complexity to memory, we want to not simply understand individual models, but understand the space of all possible models within this family.

Towards a general theory of synaptic complexity



How does the structure of a synaptic dynamical system (M^{pot} and M^{dep}) determine its function, or memory curve SNR(t)?

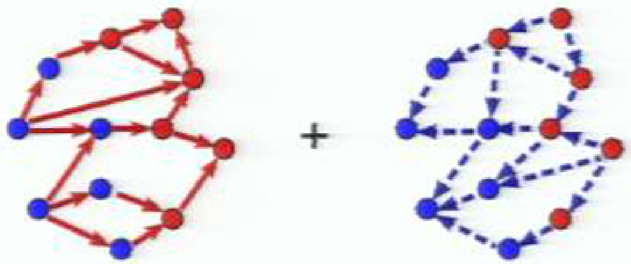
What are the fundamental limits of achievable memory over all possible choices of synaptic dynamical systems?

What is the structural organization of synaptic dynamical systems that achieve these limits?

What theoretical principles can control combinatorial explosion in the number of possible models as M increases?

Imposing a theoretical order on synaptic dynamics

As the synaptic population undergoes continuous modification, the internal state stochastically wanders around according to a forgetting process:



$M^{forget} = f^{pot} * M^{pot} + f^{dep} * M^{pot}$

This forgetting process has:

An equilibrium probability distribution of state occupancy: p_i^{∞}

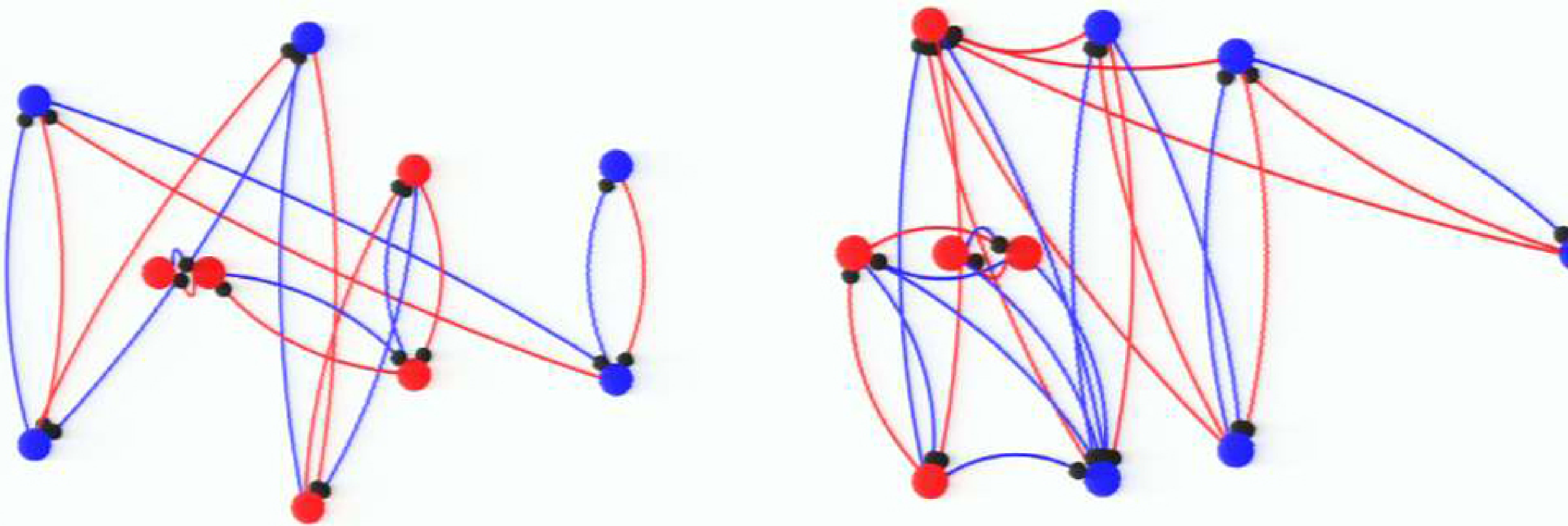
And a mean first passage time matrix from state i to j: T_{ij}

$$\eta_i^{\mathrm{pot}} \equiv \sum_{j \in \mathrm{pot}} T_{ij} \, p_j^{\infty}$$

Starting from state i, the average time it takes to get to the potentiated states, weighted by their equilibrium probability.

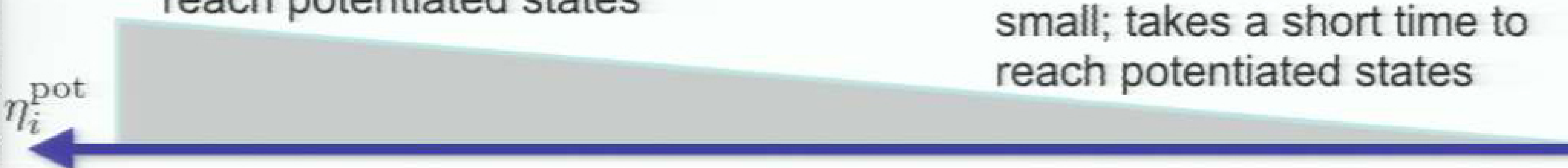
Order states from left to right in order of decreasing η_i^{pot}

Topological ordering from first passage times

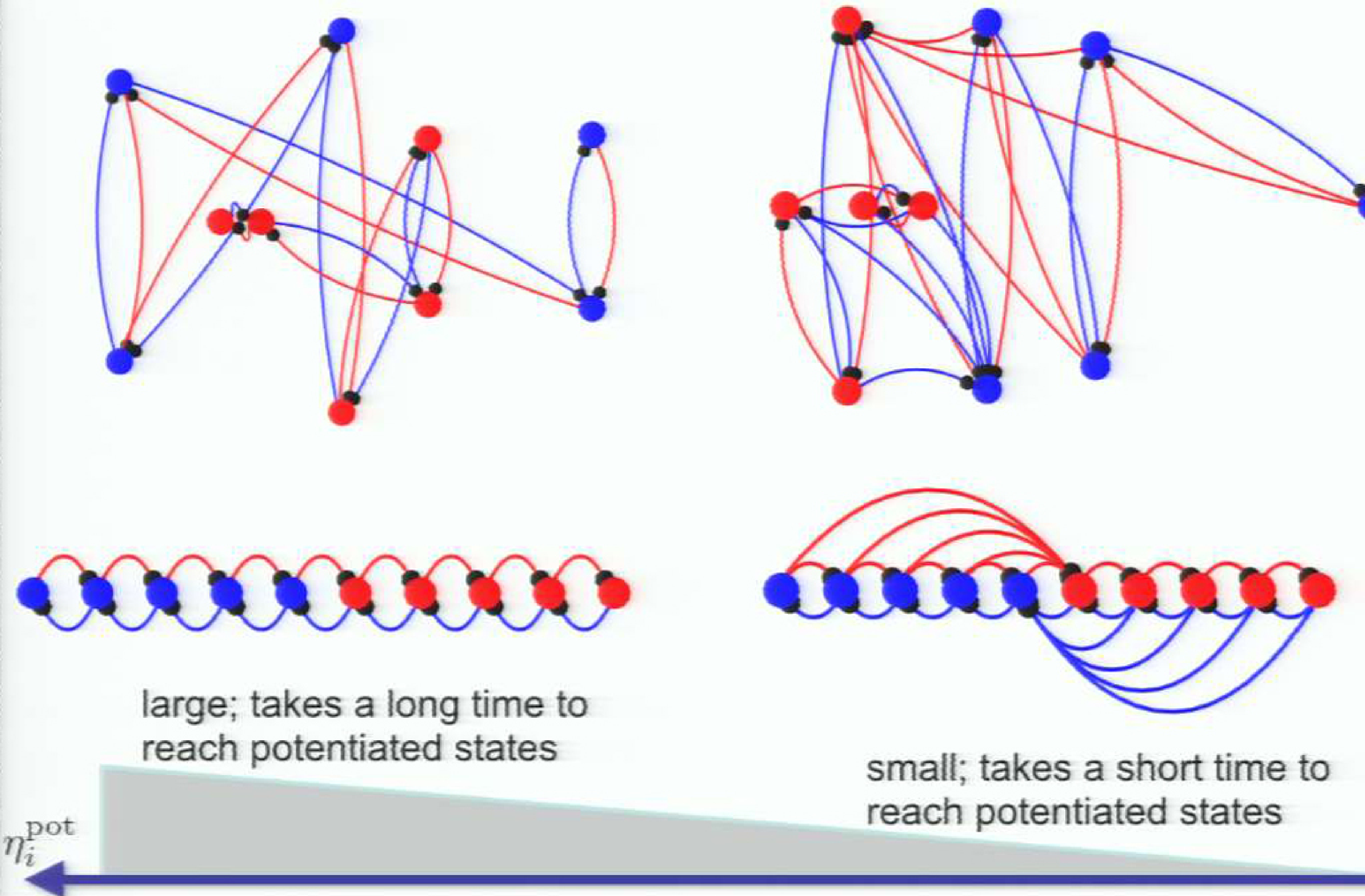


large; takes a long time to reach potentiated states

small; takes a short time to reach potentiated states



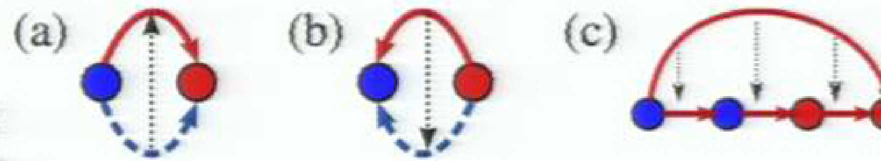
Topological ordering from first passage times



Optimal synapses have a simple structure in this order

Consider optimizing the area under the memory curve:

When states are placed in this order,



(a) M^{pot} should only go from left to right

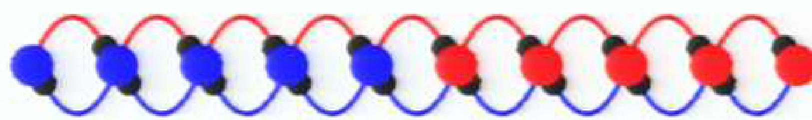
(b) M^{dep} should only go from right to left

(c) We can remove shortcuts in both M^{pot} and M^{dep} while

(1) preserving the order

(2) preserving the equilibrium distribution

(3) increasing the area



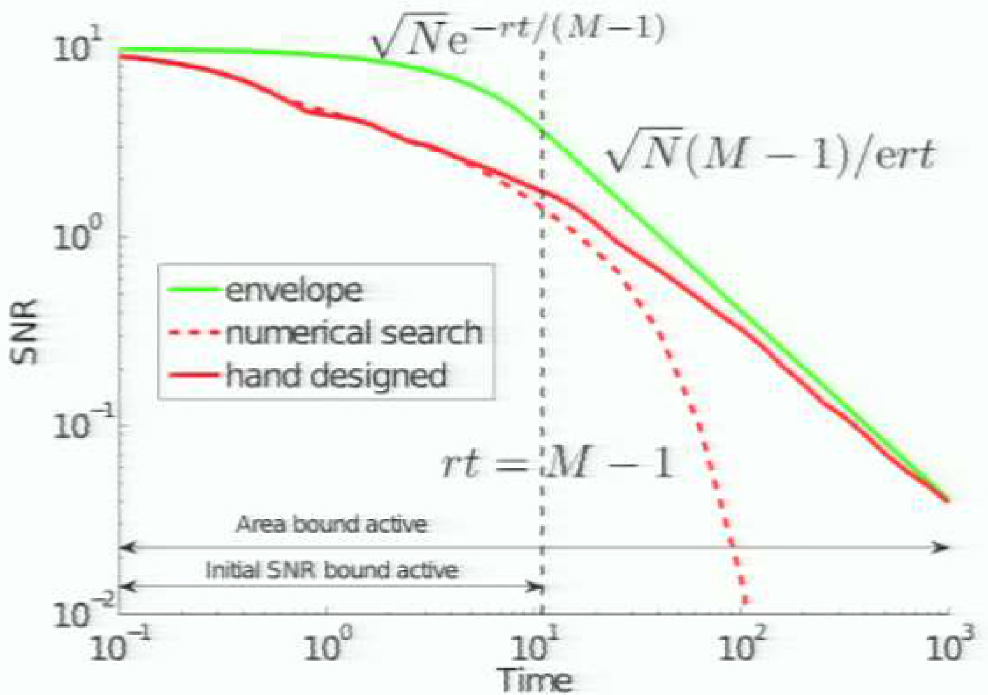
⇒ The area under the memory curve of any synaptic dynamical system is bounded by that of a chain with the same equilibrium distribution.

Also, we show that the area of a chain cannot exceed $O(N^{1/2} M)$ for any choic of transition rates along the chain.

⇒ The area under the memory curve of any synaptic dynamical system can never exceed $O(N^{1/2} M)$.

A frontier beyond whose bourn no curve can cross

Area bound implies a maximal achievable memory at any finite time given N synapses with M internal states:



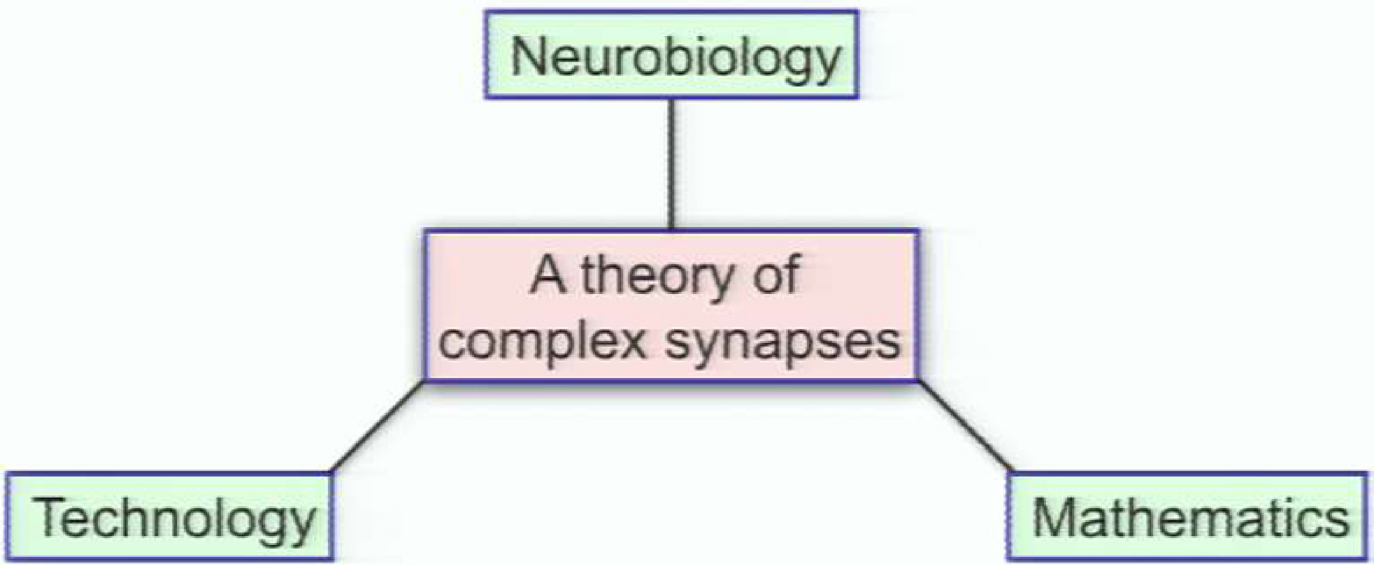
Chains with different transition rates come close to the frontier at late times.

Various measures of memory (area, frontier, lifetime) grow linearly with the number of internal states M, but grow only as the square root of the number of synapses N.

The dividends of understanding synaptic complexity

(Under review: cerebellar learning with complex synapses)

A framework for interpreting
molecular neurobiology data



The next generation of
artificial neural networks?

(Spatiotemporal credit assignment)
(Learning as message passing)

New theorems about
perturbations
to stochastic processes.

(Tighter bounds)

Acknowledgements

Subhaneil Lahiri

a.k.a.

“Subhy”

Poster F50 Tonight

Interesting conversations:

Larry Abbott
Stefano Fusi
Marcus Bena
Jascha Sohl-Dickstein
David Sussillo

Funding

National Science Foundation

Bio-X Neuroventures

Burroughs-Wellcome

DARPA

Genentech

The Alfred P. Sloan Foundation

Stanford Mind, Brain, and Computation

The Swartz Foundation

NIPS Thanks Its Sponsors

Air Force Office of Scientific Research · United States Air Force

amazon.com

Microsoft Research

Google

facebook

SKYTREE
THE MACHINE LEARNING COMPANY

TWO SIGMA

United Technologies Research Center

D E Shaw & Co

xerox

YAHOO!
LABS

IBM
Research

TOYOTA

DRW TRADING GROUP

millionshort

criteo

PDT PARTNERS

Springer
Machine Learning Journal

Disney Research