Recurrent Neural Networks as a Model to Probe Neuronal Timescales Specific to Working Memory

Robert Kim Terrence J. Sejnowski

Computational Neurobiology Laboratory, Salk Institute
University of California, San Diego (UCSD)

NeurIPS Workshop

December 14, 2019

Motivation

- Neurons important for working memory (WM) have stable and long neuronal timescales

- Neurons important for working memory (WM) have stable and long neuronal timescales
- Experimentally challenging to probe circuitry and connectivity mechanisms

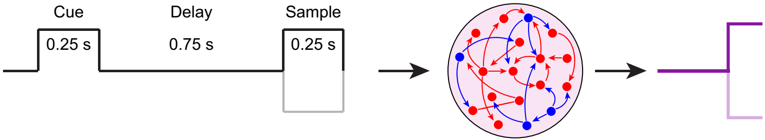
Motivation

- Neurons important for working memory (WM) have stable and long neuronal timescales
- Experimentally challenging to probe circuitry and connectivity mechanisms
- Recurrent neural network (RNN) model can be useful

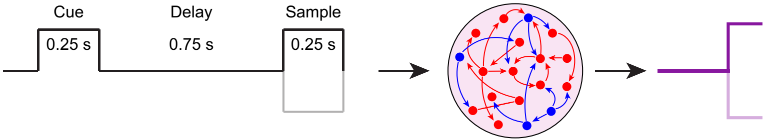
Motivation

- Neurons important for working memory (WM) have stable and long neuronal timescales
- Experimentally challenging to probe circuitry and connectivity mechanisms
- Recurrent neural network (RNN) model can be useful

Circuit/network mechanisms required for stable temporal receptive fields critical for WM maintenance

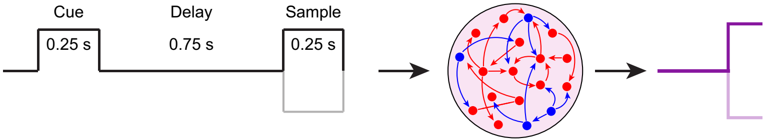


- **Continuous RNNs** converted to **Leaky Integrate-and-Fire (LIF) RNNs** [1]



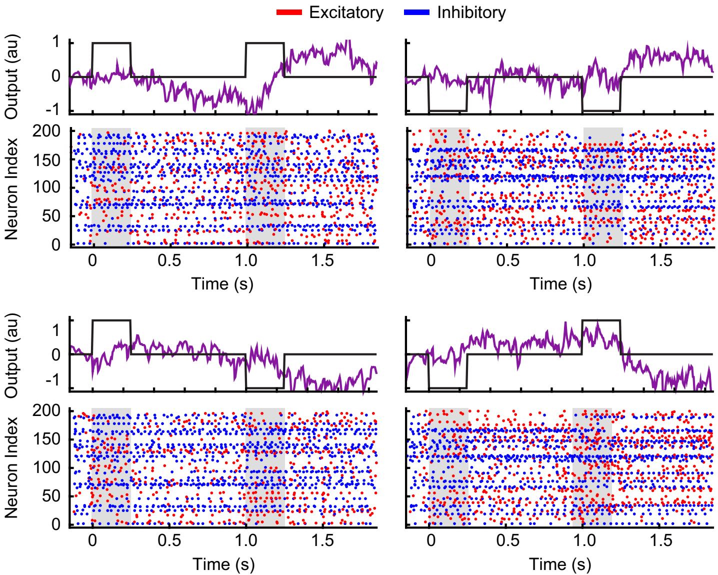
- **Continuous RNNs** converted to **Leaky Integrate-and-Fire (LIF) RNNs** [1]
- 40 RNNs ($N = 200$) trained to perform a delayed match-to-sample (DMS) task

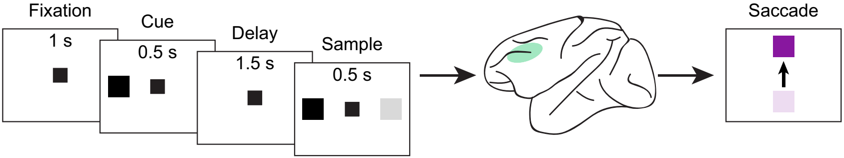
Spiking RNN Model



- **Continuous RNNs** converted to **Leaky Integrate-and-Fire (LIF) RNNs** [1]
- 40 RNNs ($N = 200$) trained to perform a delayed match-to-sample (DMS) task
- 80% excitatory and 20% inhibitory

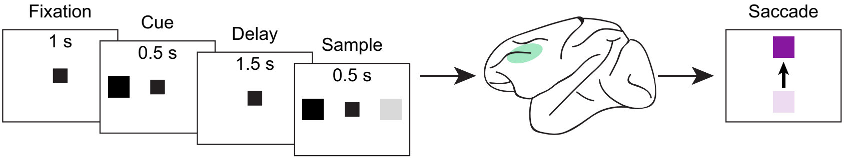
Spiking RNN Model





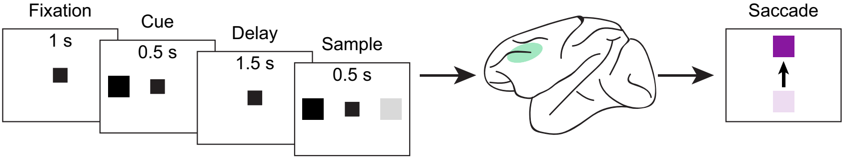
- Public dataset (`crcns.org`) – Constantinidis lab [2–4]

Experimental Data



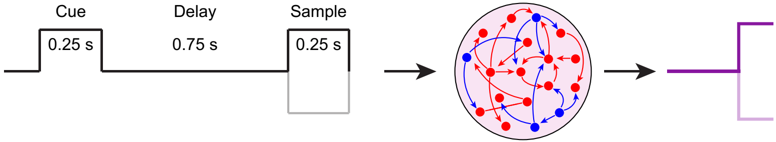
- Public dataset (`crcns.org`) – Constantinidis lab [2–4]
- **Four monkeys** trained to perform two delayed match-to-sample tasks: spatial and feature tasks

Experimental Data

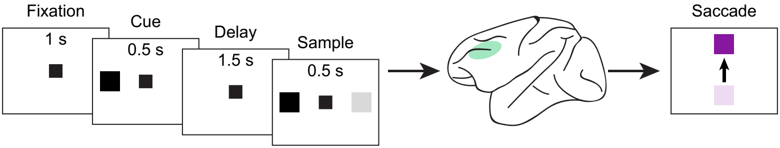


- Public dataset (`crcns.org`) – Constantinidis lab [2–4]
- **Four monkeys** trained to perform two delayed match-to-sample tasks: spatial and feature tasks
- **959 dorsolateral prefrontal cortex** (dlPFC) units

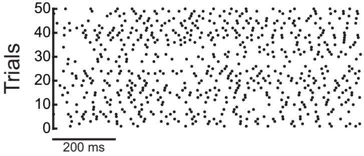
Spiking RNN Model



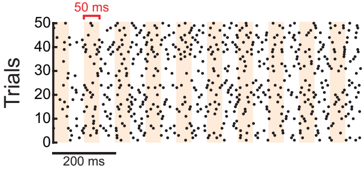
Experimental Data



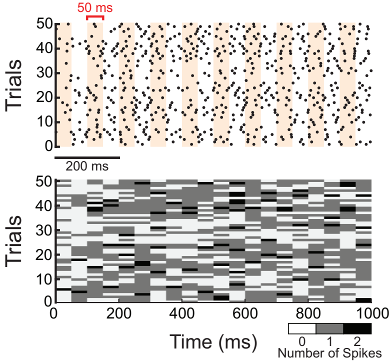
- Spike-count autocorrelation during **fixation**



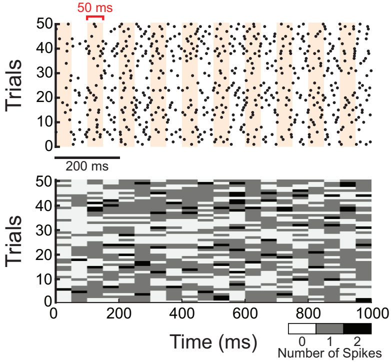
- Spike-count autocorrelation during **fixation**
- Spike counts in successive time bins ($w = 50$ ms)

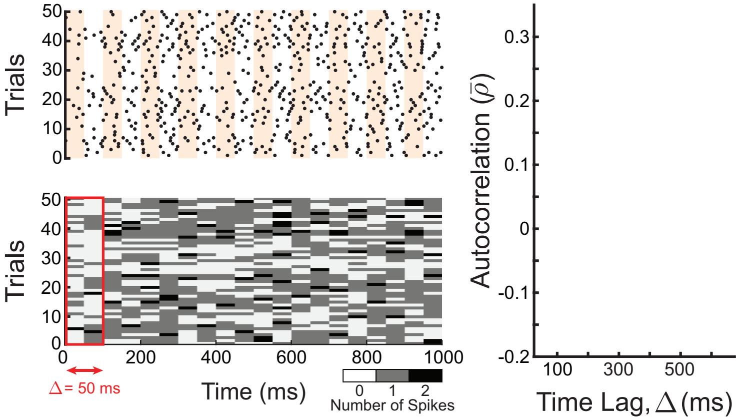


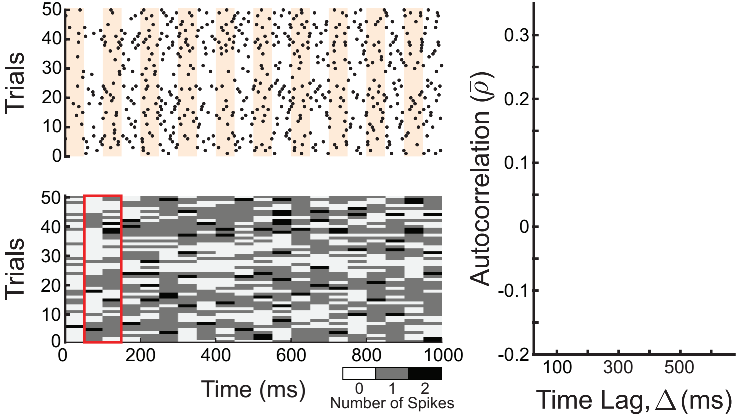
- Spike-count autocorrelation during **fixation**
- Spike counts in successive time bins ($w = 50$ ms)

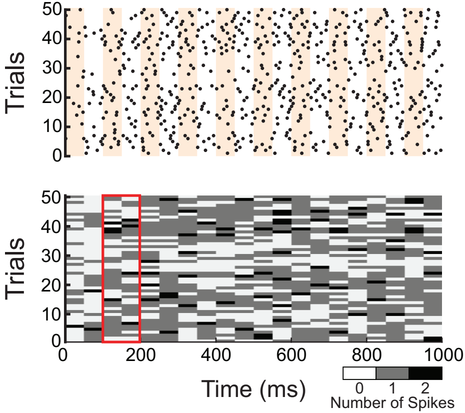


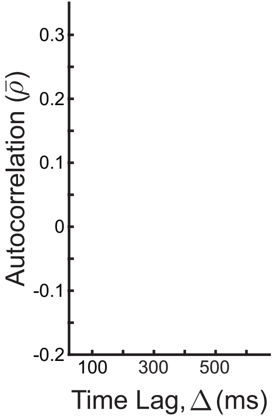
- Spike-count autocorrelation during **fixation**
- Spike counts in successive time bins ($w = 50$ ms)
- Correlation between two time bins separated by a lag (Δ)

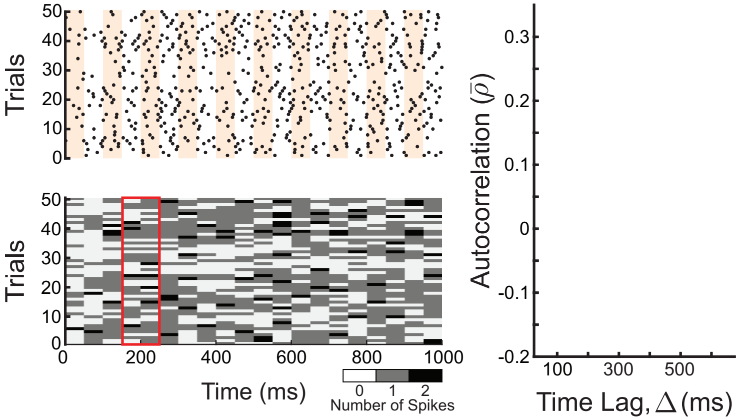


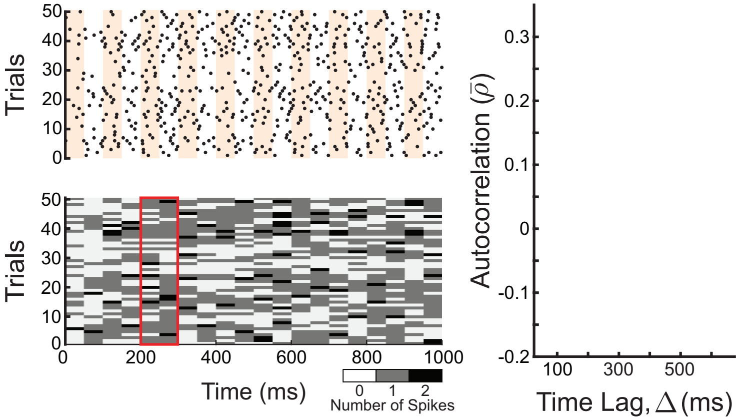

Trials
50
40
30
20
10
0
Trials
0
200
400
600
800
1000
Δ = 50 ms
Time (ms)
0 1 2
Number of Spikes
Autocorrelation ($\bar{\rho}$)
0.3
0.2
0.1
0
-0.1
-0.2
100
300
500
Time Lag, Δ (ms)

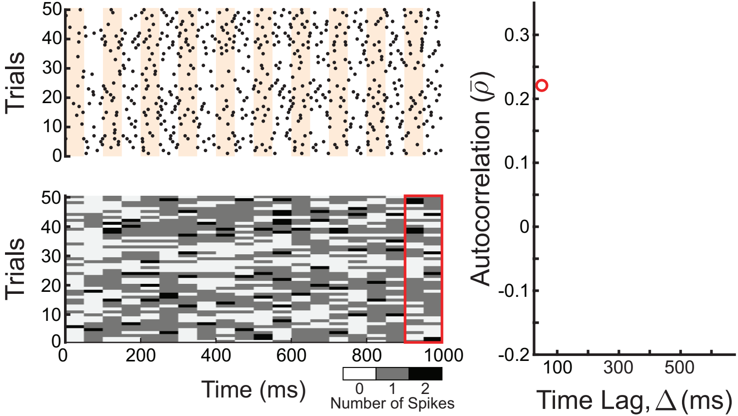

Trials
50
40
30
20
10
0
Trials
50
40
30
20
10
0
0
200
400
600
800
1000
Time (ms)
0
1
2
Number of Spikes
Autocorrelation ($\bar{\rho}$)
0.3
0.2
0.1
0
-0.1
-0.2
100
300
500
Time Lag, Δ (ms)


Trials
Trials
0
200
400
600
800
1000
Time (ms)
0 1 2
Number of Spikes

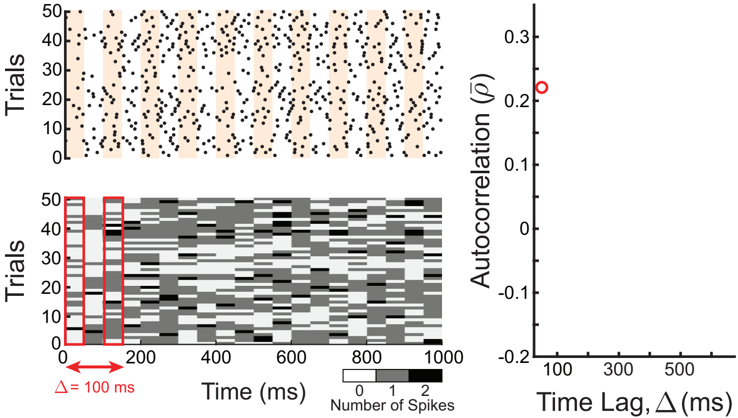

Autocorrelation ($\bar{\rho}$)
Time Lag, Δ (ms)


Trials
50
40
30
20
10
0
Trials
50
40
30
20
10
0
0
200
400
600
800
1000
Time (ms)
0
1
2
Number of Spikes
Autocorrelation ($\bar{\rho}$)
0.3
0.2
0.1
0
-0.1
-0.2
100
300
500
Time Lag, Δ (ms)

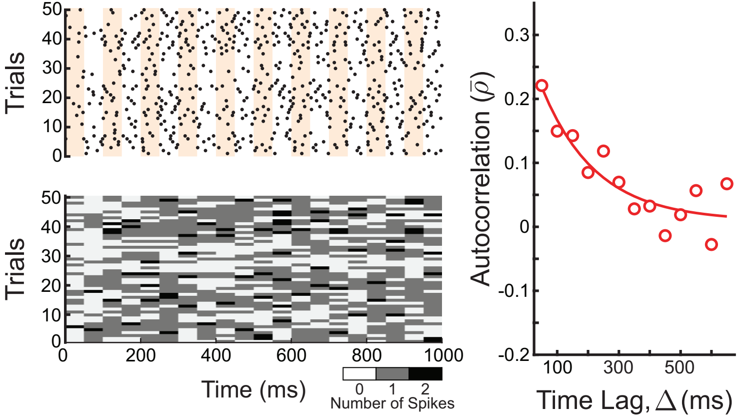

Trials
50
40
30
20
10
0
Trials
50
40
30
20
10
0
0
200
400
600
800
1000
Time (ms)
0 1 2
Number of Spikes
Autocorrelation ($\bar{\rho}$)
0.3
0.2
0.1
0
-0.1
-0.2
100
300
500
Time Lag, Δ (ms)



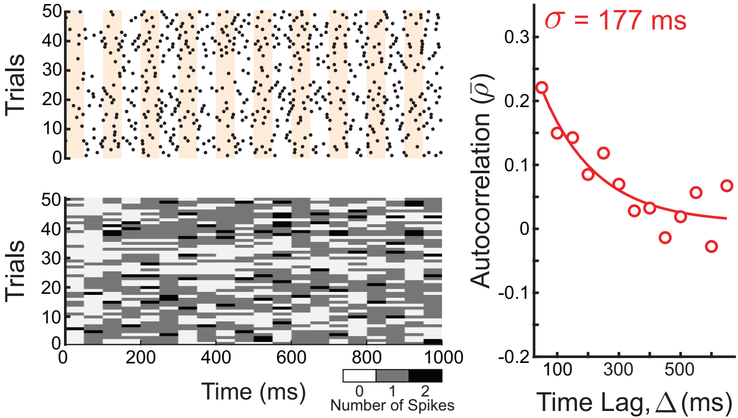
Trials
50
40
30
20
10
0
Trials
50
40
30
20
10
0
0
200
400
600
800
1000
Time (ms)
0
1
2
Number of Spikes
Autocorrelation ($\bar{\rho}$)
0.3
0.2
0.1
0
-0.1
-0.2
100
300
500
Time Lag, Δ (ms)


Trials
50
40
30
20
10
0
Trials
50
40
30
20
10
0
0
200
400
600
800
1000
Δ = 100 ms
Time (ms)
0 1 2
Number of Spikes
Autocorrelation ($\bar{\rho}$)
0.3
0.2
0.1
0
-0.1
-0.2
100
300
500
Time Lag, Δ (ms)

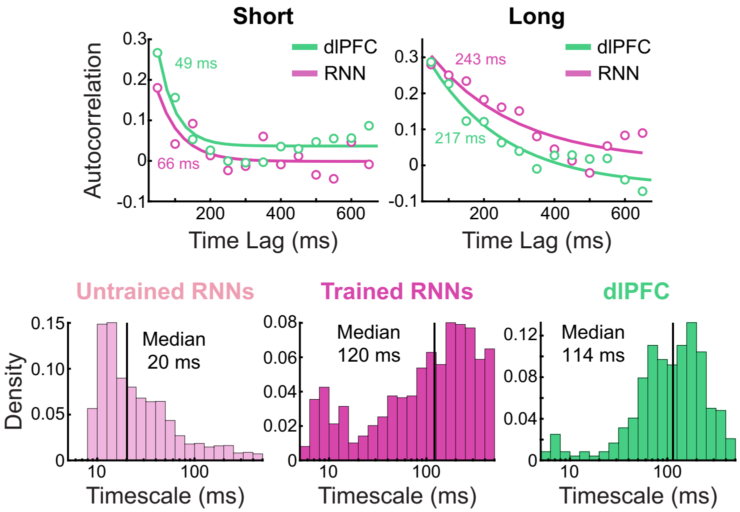
Neuronal Timescales



$$\bar{\rho}(\Delta) = A\left(\exp\left(-\frac{\Delta}{\sigma}\right) + B\right)$$

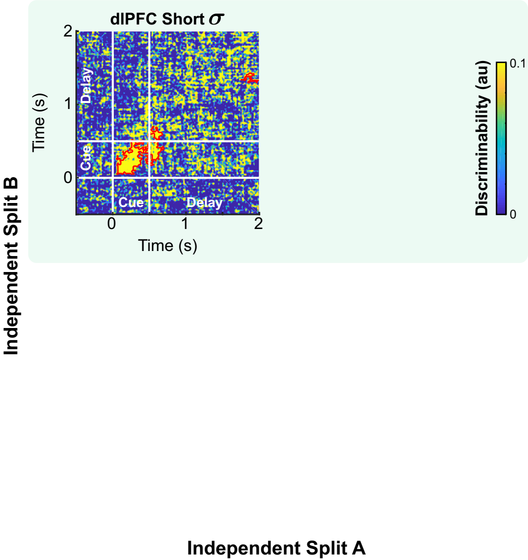


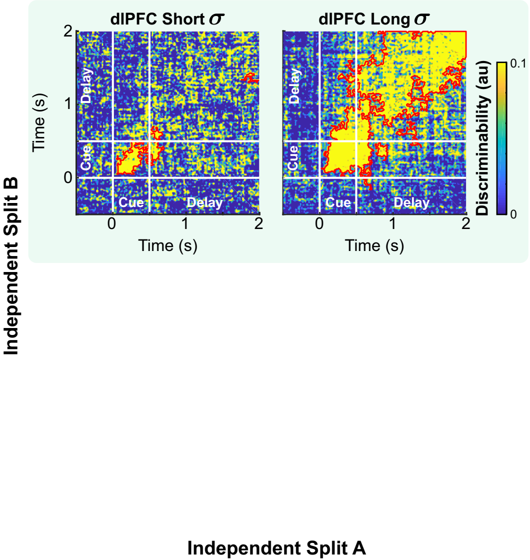
$$\bar{\rho}(\Delta) = A\left(\exp\left(-\frac{\Delta}{\sigma}\right) + B\right)$$



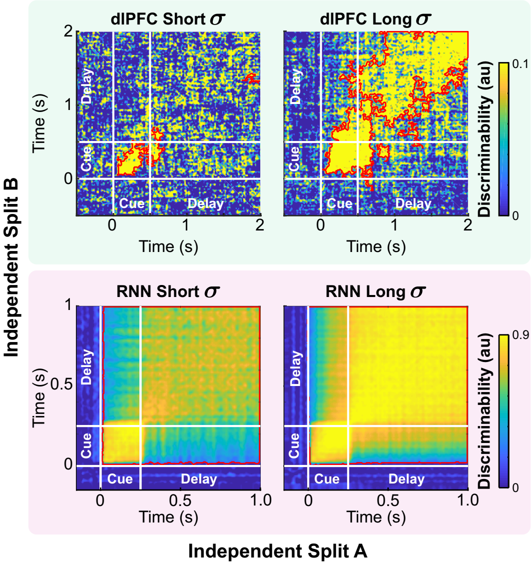
Short
Long
Autocorrelation
0.3
0.2
0.1
0
-0.1
200
400
600
Time Lag (ms)
dlPFC
RNN
49 ms
66 ms
243 ms
217 ms
Untrained RNNs
Trained RNNs
dlPFC
Density
Median
20 ms
Median
120 ms
Median
114 ms
0.15
0.10
0.05
0.08
0.06
0.04
0.02
0.12
10
100
Timescale (ms)

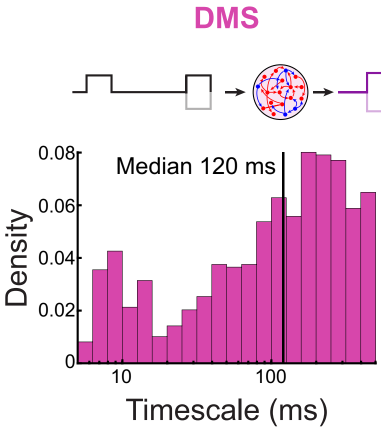
Long σ units involved with stable coding

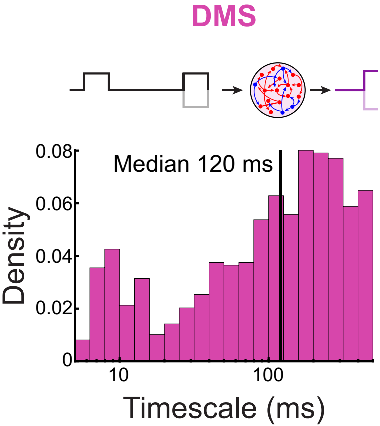


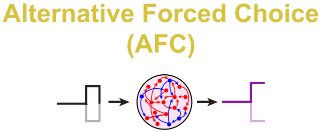

dlPFC Short σ
dlPFC Long σ
Time (s)
Delay
Cue
Cue
Delay
0
1
2
Discriminability (au)
0.1
0
Independent Split B
Independent Split A

Long σ units involved with stable coding

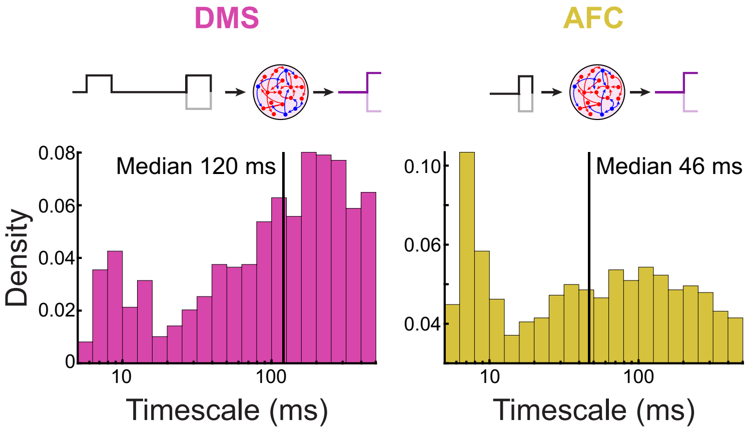


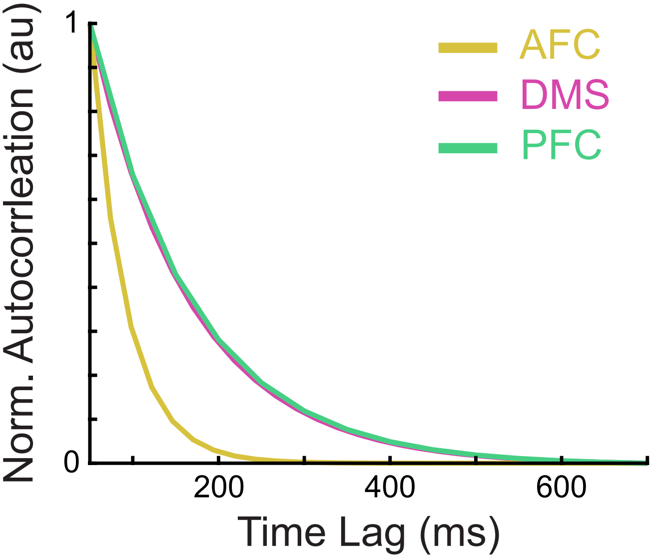

DMS
Median 120 ms
0.08
0.06
0.04
0.02
0
Density
10
100
Timescale (ms)


DMS
0.08
0.06
0.04
0.02
0
Median 120 ms
Density
10
100
Timescale (ms)


Alternative Forced Choice
(AFC)

WM-specific Neuronal Timescales




Norm. Autocorrleation (au)
1
0
200
400
600
Time Lag (ms)
AFC
DMS
PFC

- Striking similarities b/w RNN model of WM and experimental data

Summary

- Striking similarities b/w RNN model of WM and experimental data
- Both utilize units with stable temporal receptive fields to perform WM

Summary

- Striking similarities b/w RNN model of WM and experimental data
- Both utilize units with stable temporal receptive fields to perform WM
- Need to characterize network/circuit dynamics that lead to long timescales

Thank you!

1. Kim, R., Li, Y. & Sejnowski, T. J. Simple framework for constructing functional spiking recurrent neural networks. *Proceedings of the National Academy of Sciences* **116,** 22811–22820 (2019).
2. Qi, X.-L., Meyer, T., Stanford, T. R. & Constantinidis, C. Changes in Prefrontal Neuronal Activity after Learning to Perform a Spatial Working Memory Task. *Cerebral Cortex* **21,** 2722–2732 (2011).
3. Meyer, T., Qi, X.-L., Stanford, T. R. & Constantinidis, C. Stimulus Selectivity in Dorsal and Ventral Prefrontal Cortex after Training in Working Memory Tasks. *Journal of Neuroscience* **31,** 6266–6276 (2011).
4. Constantinidis, C., Qi, X.-L. & Meyer, T. Single-neuron spike train recordings from macaque prefrontal cortex during a visual working memory task before Single-neuron spike train recordings from macaque prefrontal cortex during a visual working memory task before and after training. *CRCNS.org* (2016).